Volume IV
Issue II
Autumn 2019

Editor-In-Chief
Jared Kang

Proudly Supported By
Cambridge University Law Society

CAMBRIDGE LAW REVIEW

Volume IV
Issue ii
Autumn 2019

Editor-In-Chief

Jared Kang

Proudly Supported By

Cambridge University Law Society

Cambridge Law Review

Honorary Board
The Right Honourable The Lord Millett PC
Judge Hisashi Owada
Judge Awn Al Khasawneh
The Right Honourable Sir John Laws PC
The Honourable Sir Jeremy Cooke QC
Justice Anselmo Reyes SC
Professor Malcolm Shaw QC
Michael Blair QC
Jern-Fei Ng QC

Editor-In-Chief
Jared Kang

Vice Editors-In-Chief
Neli Frost
Rabin Kok

Managing Editor
Hope Williams

Managing Editor
Hope Williams
Roystan Ang Jun Teck

Senior Editors
Benjamin Newman
Chloe Francis
Darius Ng
Despoina Georgiou
Nicole Chia

Editors
Annie Mackley
Ava Mayer
Cherie Ho
Emilia Howard
Ernest Goh Jun Hong
Felipe Antonio Armijo Losilla
Louis Cheng
Oliver Hailes
Ran Xu
Tatiana Kurschner
Vanessa Cheok
Wesley Poh
Zihang Liu

International Editors
Alex Yeung
The University of Hong Kong
Brendan Clift
University of Melbourne
Daryl Pang
National University of Singapore, University of Toronto
Georgina Dimopoulos
University of Melbourne
Jenny Poon
University of Western Ontario
Ong Kye Jing
National University of Singapore
Tanaya Sanjay Rajwade
National Law University, Delhi

Table of Contents

Criminal Jurisdiction over the Internet: Jurisdictional Links in the Cyber Era 1
Dominik Zając

Regulating Virtual Currency Payment Systems 29
Sijuade Animashaun

Old is Sometimes Better: The Case For Using Existing Law To Face The Challenges Of The Digital Age 68
Riccardo de Caria

The Enabler Theory And Atrocity Crimes 83
Shahram Dana

In The Light Of Different National Circumstances: Equity Under The Paris Agreement 105
Elkanah Oluwapelumi Babatunde

Sumangali System: Is The Truth Ugly? 138
Yamuna Menon

Justice, Jurisdiction, And R (On The Application Of Privacy International)
Owen Alexander Sparkes 148

Vedanta Resources V Lungowe: A Pre-Existing Pocket Of Negligence, Or A Novel Scenario? 176
Tan Jin Hsi, Gabriel

Personal Injury, Autonomy, and Johnson Matthey 185
Brian Ip

Cambridge Law Review (2019) Vol IV, Issue ii, 1–28

Criminal Jurisdiction over the Internet: Jurisdictional Links in the Cyber Era

Dominik Zając*

I. Introduction

A result of the creation of the Internet is that the real world no longer is the only space in which interpersonal interaction occurs.[1] Now, a completely new, somehow parallel-to-reality plane exists that escapes geographical limitations. It forces redefinition of the concepts of sovereignty and jurisdiction.[2] Individual countries, to some extent grouped, have begun to seek any foothold that allows them to regulate and punish behaviours undertaken by Internet users.

This study is a critical analysis of solutions used to determine the scope of criminal jurisdiction in cyberspace. Considering the intensive development of social interactions undertaken using the Internet, it appears justified to move away from a rigid model of jurisdictional rules and shift to a discursive model based on weighing the interests of states.

The considerations are divided into six parts. Section II presents a short description of a method based on classical jurisdictional rules. Section III includes

* Dominik Zając, Ph.D., is an Assistant Professor at the Department of Criminal Law of The Jagiellonian University in Krakow. This article was written with the support of the Foundation for Polish Science (FNP).

1 See *inter alia*: Stine Gotved, "Time and space in cyber social reality" (2006) 8 New Media & Society 467–86; Toni C Antonucci, Kristine J Ajrouch and Jasmine A Manalel, "Social Relations and Technology: Continuity, Context, and Change" (2017) 3 Innov Aging 1–9.

2 Michael N Schmitt, Liis Vihul (eds), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (CUP 2017) 12; Dan L Burk, "Muddy Rules for Cyberspace" (1999) 21 Cardozo L Rev 121, 122; Michael N Schmitt, Liis Vihul, "Respect for Sovereignty in Space" (2017) 95 Tex L Rev 1639; Joanna Kulesza, *International Internet Law* (Roudlegde 2012) 2–3.

a discussion of traditional approach defects, particularly considering those features that prevent the application of that traditional approach to behaviours undertaken using cyberspace. Section IV discusses an alternative method of determining scope of jurisdiction based on important elements of the social situation (jurisdictional links). Sections V and VI parts are devoted to an analysis of individual nexuses. In Section VII, the method of weighing the significance of the links is presented. This method allows granting a particular state the right to regulate or impose a penalty for a given behaviour on the Internet.

II. Principles of jurisdiction as a traditional method of establishing jurisdiction in criminal cases

The jurisdiction principles are currently the substantive basis for making claims for regulation or penalisation.[3] To demonstrate the existence of the power to legislate and enforce the law, the state relies on one of the four main jurisdictional principles: territoriality,[4] the active personality principle,[5] the passive personality principle[6] and the protective principle.[7] This catalogue is broadened in the case of competences that constitute *ius puniendi*,[8] in which the authorisations are also based on the principle of vicarious jurisdiction[9] and the rule of universal jurisdiction.[10] The above are confirmed by argumentation conducted in the context of disputes

[3] Adria Allen, "Internet Jurisdiction Today" (2001) 22 Nw J Int'l L & Bus 69, 75; Ray August, "International Cyber-Jurisdiction: A Comparative Analysis", (2002) 39 ABLJ 531, 534; Christopher Kuner, "Data Protection Law and International Jurisdiction on the Internet (Part I)" (2010) 18 International Journal of Law and Information Technology 176, 188–191; Kulesza (n 2) 6.

[4] Cherif M Bassiouni, *Crimes Against Humanity: Historical Evolution and Contemporary Application* (CUP 2011) 279; *"Extraterritorial criminal jurisdiction. Council of Europe"* European Committee on Crime Problems (1992) 3 Criminal Law Forum 441, 446.

[5] Cedric Ryngaert, *Jurisdiction in International Law* (2nd Ed, OUP 2015) 89; Bassiouni (n 4) 279; Danielle Ireland-Piper, "Prosecutions of Extraterritorial Criminal Conduct and the Abuse of Rights Doctrine" (2013) 9 Utrecht L Rev 68, 74; William R Slomanson, *Fundamental Perspectives on International Law* (Wadsworth Publishing 2010) 250.

[6] Ryngaert (n 5) 93; Bassiouni (n 4) 279; Geoffrey R Watson, "The Passive Personality Principle" (1993) 28 Tex Int'l LJ 1 18; Jonathan O Hafen, "International Extradition: Issues Arising under the Dual Criminality Requirement" (1992) BYU L Rev 19, 218; John G McCarthy, "The Passive Personality Principle and Its Use in Combatting International Terrorism" (1989) 13 Fordham Int'l LJ 298, 301.

[7] Ryngaert (n 5) 97; Bassiouni (n 4) 279; *Extraterritorial* (n 4) 451.

[8] Kai Ambos, "Punishment without a Sovereign? The Ius Puniendi Issue of International Criminal Law: A First Contribution towards a Consistent Theory of International Criminal Law" (2013) 33 OJLS 293, 297–8; Andrzej Sakowicz, *Zasada ne bis in idem w prawie karnym* (Temida2 2011) 138.

[9] Ryngaert (n 5) 103; *Extraterritorial* (n 4) 452.

[10] Ryngaert (n 5) 106; *The Princeton Project on Universal Jurisdiction (Program in Law and Public Affairs Princeton University 2001)* 28; Hafen (n 6) 219; Bassiouni (n 4) 280.

over competences between countries, in which individual parties derive their rights from the facts of the application of these principles.[11] Recourse to a given principle is made to justify the right to exercise state authority.

All jurisdiction principles mentioned above have evolved in the course of the historical development of international law.[12] They are, in fact, of a customary nature. They refer to the bonds existing between a specific social situation and the state. Over the centuries, however, there has been a specific looping in this area. Its effects are particularly strongly felt in today's realities. Historical analysis permits advancing the thesis that the rules of jurisdiction have developed as a legal description of socially significant connections and have slowly "ossified". As social relationships evolved and new connections developed that were of great importance, no new jurisdictional rules were created; instead, there was an attempt to enter such relationships and connections into already existing content. This approach also has been applied to social situations that occur in cyberspace.

Particularly noteworthy in this context are two circumstances, the recognition of which is essential for describing the weaknesses of the classical jurisdictional model.

First, the jurisdiction rules were included in the national legal systems at a time when virtually all human activities were of a physical nature. Social situations that were regulated by law were clearly defined in space.[13] The resulting consequences were characterised by a small spatial scope, which in principle was easy to predict from an ex-ante perspective. This point is evidenced by academic examples used to discuss cross-border crimes, in terms of elements such as causing

[11] Ellen S Podgor, "Cybercrime: Discretionalnary Jurisdiction" (2008–2009) 47 U Louisville L Rev 727, 729.

[12] Howard J Grootes, "Territorial Jurisdiction in Cyberspace" (2002) 4 Or Rev Int'l L 17, 28; *Tallin* (n 2) 51.

[13] Georgios I Zekos, "State Cyberspace Jurisdiction and Personal Cyberspace Jurisdiction" (2007) 15 International Journal of Law and Information Technology, 1, 20.

a result by an archery shot.[14] Cross-border was a marginal problem; therefore, it did not require the development of sophisticated, dogmatic instruments.

Second, the system of analysed principles was developed many years before the modern concept of human rights was verbalised.[15] Consequently, the former does not consider the rights of the individual, which directly affects how the limits of the state's jurisdiction are determined.[16] The jurisdictional principles focus on inter-state relations. The reference point here is the need for protection of the sovereignty of independent entities of international law, which has important consequences; an inter-state dispute can always be resolved ex post at the political level.[17] Such a situation is unacceptable when considering an individual's right to become familiar with the content of the law in force.

III. Defectiveness of the traditional method based on jurisdictional principles

Determining the limits of state jurisdiction based on the system of existing jurisdictional rules is currently counter-effective. Patching a leaky system by means of *ex post* political arrangements does not solve the problems that 21st century

[14] Eduard Treppoz, "Jurisdiction in Cyberspace" (2016) 26 Swiss Rev Int'l E L 273, 275.

[15] *Tallin* (n 2) 179; Wojciech Burek, *Zastrzeżenia do traktatów z dziedziny praw człowieka* (Instytut Wydawniczy EuroPrawo 2012) 45; Anne Clunan, "Redefining Sovereignty: Humanitarianism's Challenge to Sovereign Immunity" in Noha Shawki and Meacheline Cox (eds), *Negotiating Sovereignty and Human Rights: Actors and Issues in Contemporary Human Rights Politics* (Padstow 2009) 7–27; Jean L Cohen, *Globalization and Sovereignty: Rethinking Legality, Legitimacy, and Constitutionalism* (CUP 2012) 159–78; Dominik Zając, *Odpowiedzialność karna za czyny popełnione za granicą* (KIPK and Wolters Kluwer 2017) 246; Oona A Hathaway, "International Delegation and State Sovereignty" (2008) 71 Law & Contemp Probs 115, 145–8; Stephen D Krasner, *Sovereignty: Organized Hypocrisy* (Princeton University Press 1999) 125; Robert Jackson, *Sovereignty: The Evolution of an Idea* (CUP 2007) 114–34. The above trend was reflected in the case law of the ICTY, in which it was noted: „the impetuous development and propagation in the international community of human rights doctrines, particularly after the adoption of the Universal Declaration of Human Rights in 1948, has brought about significant changes in international law, notably in the approach to problems besetting the world community. A State-sovereignty-oriented approach has been gradually supplanted by a human-being-oriented approach. Gradually the maxim of Roman law *hominum causa omne jus constitutum est* (all law is created for the benefit of human beings) has gained a firm foothold in the international community as well." – *Prosecutor v Dusko Tadic a/k/a "Dule", Decision on the Defence Motion or Interlocutory Appeal on Jurisdiction, Decision of 2 October 1995,* <http://www.icty.org/x/cases/tadic/acdec/en/51002.htm> (accessed 5 November 2018).

[16] Greg Y Sato, "Should Congress Regulate Cyberspace" (1997) 20 Hastings Comm & Ent LJ 699, 716–7.

[17] Jedynie jako przykład wskazac należy tutaj na sprawe Cuttinga – see John B Moore, *Report on Extraterritorial Crime and the Cutting Case* (United States Department of State 1887).

societies face in this respect.[18] Two major defects exist in the traditional system. First, recourse to jurisdictional rules does not allow for a categorical *ex ante* determination, *i.e.*, at the moment of making decisions about a given behaviour, of the content of the norm binding the perpetrator. Second, it is not possible to sensibly weigh these principles and, consequently, to determine the content of the law in force.

A. Lack of recognisability of the jurisdictional basis at the moment of action or omission of individual

Traditional jurisdictional rules describe the powers of the state without a consistent separation between prospective and responsive competences.[19] The first group of competences decide on the possibility of shaping future behaviour of perpetrators.[20] The second group allow punishing the individual in the case of breaking the law, which regulates his or her behaviour.[21]

The first group includes the competence to set regulatory norms.[22] Thus, the state influences how people behave, indicating, for example, that one should move to the right side of the road and should not download illegal software from the network. This competence is prospective because it is not an answer to a past event but rather is supposed to model future behaviour. The second group includes the right to punish. It has a responsive character and allows only a subsequent reaction to a past event (crime). For the modelling of future behaviours to be possible, it must be possible to determine the content of the standard binding the entity at the moment of action or omission.[23] The criterion of the application of a norm in space must therefore be based on an element of social situations that is recognisable *ex ante*.

Not all of the proposed rules meet the above criteria. Some of them refer, for example, to the place in which the effects of behaviour or financial gain occur.[24]

[18] Edward Lee, "Rules and Standards for Cyberspace" (2002) 77 Notre Dame L Rev 1275, 1279, 1281; Kulesza (n 2) 30; Zekos (n 13) 15; Jennifer Daskal, "Borders and Bits" (2018) 79 Vand L Rev 179, 222; Jennifer Daskal, "The Un-Territoriality of Data" (2015) 124 Yale L J 326, 330; Uta Kohl, *Jurisdiction and the Internet* (CUP 2007) 59.

[19] *Extraterritorial* (n 4) 458. The document indicates the following: "Legislative jurisdiction and judicial jurisdiction coincide in the case of criminal law. The national courts apply, in principle even with respect to offences which may have been committed outside national territory".

[20] Edward Lee (n 18) 1314; C Alchourron, E Bulygin, *Normative Systems,* (Springer 1971) 42.

[21] Hans Kelsen, "The Pure Theory of Law and Analytical Jurisprudence" in *What is Justice? Justice, Law and Politics in the Mirror of Science* (University of California Press 1971) (1957) 275.

[22] Zhang Xinbao, Xu Ke, "A Study on Cyberspace Sovereignty" (2016) 4 China Legal Sci 33, 59.

[23] Kohl (n 18) 116.

[24] Ryngaert (n 5) 76; Hafen (n 6) 216; *Extraterritorial* (n 4) 446; Treppoz (n 14) 279.

Considering the nature of these circumstances, the perpetrator might not even be able to predict whether the effects will occur in the geographical space of another state.

The inability to recognise the content of the norm at the moment of action excludes the possibility of its use in the process of assessing the behaviour of the perpetrator. This limitation not only results from the principle of *nullum crimen sine lege*[25] but also finds a deeper practical justification. It is impossible to require a certain behaviour from someone but at the same time not give him a chance to recognise what the behaviour should be. Therefore, at the level of defining the scope of competence for standardisation, it is necessary to reject all of those jurisdictional principles that are based on future and uncertain circumstances.

B. Lack of possibility to weigh the principles of jurisdiction

Even when, based on the realities of a given social situation, it is possible to determine which jurisdictional principle is applicable, doing so will not eliminate the conflict of jurisdiction.[26] In international law, there is no universally accepted hierarchy in this respect.[27] In the case of the Lotus tanker,[28] the Permanent Court of International Justice has unequivocally indicated that the only circumstance limiting the state's exercise of competence is the sovereignty of other states and

[25] Mohamed Shahabuddeen, "Does the Principle of Legality Stand in the Way of Progressive Development of Law?" (2004) 2 J Int'l Crim Just 1007, 1008; Franz von Liszt, "The Rationale for the Nullum Crimen Principle" (2010) 5 J Int'l Crim Just 1010.

[26] Jack L Goldsmith, "Against Cyberanarchy" (1999) 40 University of Chicago Law Occasional Paper, 1, 16.

[27] Ryngaert (n 5) 271; Florian Jessberger, W Kaleck, *Concurring Criminal Jurisdictions under International Law, The European Center for Constitutional and Human Rights (ECCHR)*, <https://www.ecchr.eu/fileadmin/Gutachten/Expert_Opinion_Concurrent_Jurisdictions_en_Verantwortung_Voelkerstraftaten.pdf> (accessed 5 November 2018); *General principles of international criminal law, ICRC Advisory Service on International Humanitarian Law,* <https://www.icrc.org/eng/assets/files/2014/general-principles-of-criminal-icrc-eng.pdf> (accessed 5 November 2018), 1; *Tallin* (n 2) 56.

[28] *Lotus Case, Publications of the Permanent Court of International Justice. Series A.-No. 70, September 7th, 1927, Collection Of Judgments, The Case of the S.S. LOTUS,* <https://www.icj-cij.org/files/permanent-court-of-international-justice/serie_A/A_10/30_Lotus_Arret.pdf> (accessed 5 November 2018); Schmitt and Vihul (n 2) 1650.

the binding norms of international law.[29] In practice, this ruling means that in the event of a conflict of jurisdiction, diplomatic pressure and the speed of the state's actions are crucial. There are no precise criteria that would allow us to clearly state which country has a stronger power to regulate. The victim of the above situation is first an individual who remains in a situation of uncertainty—even when it is determined which jurisdictional principles apply to his or her activity. A person who resides in the territory of Germany who downloads data protected by copyright from a site whose content is stored on servers located in the United States will thus not be able to determine which law is binding on him or her. Moreover, to demonstrate their own competence in the field of regulation and criminalisation of behaviour, both Germany and the US will invoke the principle of territoriality in its objective variant.[30]

Thus, application of the traditional jurisdictional principles does not allow for the resolution of the fundamental problems associated with the method of attributing responsibility for crimes of a cross-border nature.

IV. Jurisdictional links as arguments for the existence of state competence

In the doctrine of international law, the thesis according to which state competences are limited only by the sovereignty of other entities and the binding norms of international law is commonly accepted.[31] It is most clearly expressed in the Lotus judgment already discussed. These powers of the state are derived from sovereignty. From the perspective of the present study, the most important of them

[29] *Lotus Case* (n 28): "It does not, however, follow that international law prohibits a State from exercising jurisdiction in its own territory, in respect of any case which relates to acts which have taken place abroad, and in which it cannot rely on some permissive rule of international law. Such a view would only be tenable if international law contained a general prohibition to States to extend the application of their laws and the jurisdiction of their courts to persons, property and acts 'outside their territory, and if, as an exception to this general prohibition, it allowed States to do so in certain specific cases. But this is certainly not the case under international law as it stands at present. Far from laying down a general prohibition to the effect that States may not extend the application of their laws and the jurisdiction of their courts to persons, property and acts outside their territory, it leaves them in this respect a wide measure of discretion which is only limited in certain cases by prohibitive rules; as regards other cases, every State remains free to adopt the principles which it regards as best and most suitable". See also: Ryngaert (n 5) 22–6; Christopher Greenwood, "Sovereignty: A View from the International Bench" in Richard Rawlings, Peter Leyland and Alison Young, *Sovereignty and the Law: Domestic, European and International Perspectives* (OUP 2013) 258; Roman Kwiecień, *Teoria i filozofia prawa międzynarodowego* (Difin 2011) 115.

[30] *Extraterritorial* (n 4) 446; Hannah L Buxbaum, "Territory, Territoriality, and the Resolution of Jurisdictional Conflict" (2009) Articles by Maurer Faculty, Paper 132, 642; David G Post, "Against "Against Cyberanarchy" (2002) 17 Berkeley Tech LJ 1365, 1381; Daskal, The Un-Territoriality (n 18) 326; Kohl (n 18) 11.

[31] Zając (n 15) 73.

are the right to regulate behaviours[32] and the right to impose punishment on an individual.[33] Together, they constitute the law of punishment (*ius puniendi*).[34] The state can prohibit specific behaviour under the threat of punishment (regulatory aspect *ius puniendi*) and punish a person breaking this prohibition (repressive, procedural aspect of *ius puniendi*). The two competences are coupled together such that the imposition of punishment is possible only when the behaviour of the offender constitutes a violation of the law binding him or her at the moment and place of action or omission.

To effectively enforce the law, the state tries to show that a given social situation influences the interpersonal relationships under its protection.[35] This approach is the surest means of avoiding the accusation that its actions interfere with the sphere of exclusive rights of other states. If a given situation is connected, even non-exclusively, with the X-state, then it cannot be said that the state of Y has exclusive power over it. For this purpose, the state refers to the existence of certain elements of social situations which, from the perspective of international law, testify to the relationship that exists between them and the social situation. These elements (*e.g.*, territory and citizenship) are referred to as jurisdictional links.[36] They form the basis for validation arguments (regulatory aspect) and penalisation arguments (repressive aspect). The more important the relationship becomes the stronger will be the state's claim to set the norm or punish violation of the law.

In addition to the interests of particular states, in the process of defining the limits of spatial effectiveness of norms, it is necessary to consider the individual's rights. The individual has its own interests, the existence of which has been recognised and which are protected by the international system of human rights protection.[37] Therefore, these interests must be considered in the frames of the

[32] Willis L M Reese, "Legislative Jurisdiction" (1978) 78 Colum L Rev 1587; Austen L Parrish, "Evading Legislative Jurisdiction" (2013) 87 Notre Dame L Rev 1673, 1677; John H Knox, "A Presumption Against Extrajurisdictionality" (2010) 104 Am J Int'l L 351, 355; John H Knox, "Legislative Jurisdiction, Judicial Canons, and International Law" (2009) 100 Wake Forest Univ Legal Studies Paper No. 1349127, 2.

[33] Anthony Duff, "Responsibility, Citizenship and Criminal Law" in Anthony Duff & Suart Green (eds), *Philosophical Foundations of Criminal Law* (OUP 2011) 127.

[34] *Extraterritorial* (n 4) 456; Tomasz Ostropolski, *Zasada jurysdykcji uniwersalnej* w *prawie międzynarodowym* (Instytut Wydawczniczy EuroPrawo 2008) 20; Michał Płachta, Jurysdykcja państwa w sprawach karnych wobec cudzoziemców (1992) 111/112 Studia Prawnicze 98.

[35] Kohl (n 18) 20.

[36] Treppoz (n 14) 275; Marek Wasiński, Jurysdykcja legislacyjna państwa w prawie międzynarodowym publicznym, (2002) 673 Państwo i Prawo 56, 61; Knox, Legislative (n 32) 102; Kohl (n 18) 15.

[37] Hathaway (n 15) 146; Cohen (n 15) 178–9; Alette Smeulers and Fred Grünfeld, *International Crimes and Other Gross Human Rights Violations: A Multi- and Interdisciplinary Textbook* (Brill Nijhoff 2011) 7; Jackson (n 15) 124.

validation and penalisation discourses. The individual has the right to become acquainted with the content of the binding law. He or she has also a right to act within the limits of a compatible legal system.[38] Any ambiguity with respect to the validity of standards cannot have negative consequences for the individual. This requirement is particularly emphasised in the case of criminal law regulations. To this extent, the values protected by human rights are the basis of negative validation or penalisation arguments. They do not support the claim of any sovereign entity, but only block some of them—in the event of their being contrary to the content of human rights.

Only the joint consideration of positive and negative penalisation and validation arguments allows us to determine whether a given country has the power to regulate a given social situation or to punish a violation of law.

A. Concept of jurisdictional links and the specificity of behaviour undertaken in cyberspace

The discursive approach[39] outlined above can be successfully applied in defining the limits of state authority over the behaviour of the Internet. To this end, certain specific features of social situations that occur in cyberspace are considered.

The central point for considering the scope of jurisdiction of a state is the individual's behaviour and its consequences,[40] referred to collectively as the social situation. Such a "social situation" is a phenomenon occurring in a space-time composed of many elements. Some of them can constitute a relationship between the social situation and the state (they are referred to above as jurisdictional links). The implementation of the behaviour using the Internet does not modify the above

[38] Beth Van Schaack, "Crimen Sine Lege: Judicial Lawmaking at the Intersection of Law and Morals" (2008) 97 Geo LJ 119, 172–88; Shahram Dana, "Beyond Retroactivity to Realizing Justice: A Theory on the Principle of Legality in International Criminal Law Sentencing" (2008–2009) 99 J Crim L & Criminology 857, 867.

[39] Andrzej Grabowski, *Juristic Concept of the Validity of Statutory Law A Critique of Contemporary Legal Nonpositivism* (Springer 2013) 455.

[40] The perception of the Internet network as a separate space should be rejected—this approach does not explain anything, at the same time making it difficult to rationally assess the relationship between behaviour and the state seeking to normalize or punish. In the literature, it is rightly recommended to move away from the metaphor of "space" for Internet infrastructure. On the basis of this study, the term "cyberspace" is used only as a shorthand, describing the social situations that are undertaken with the usage of the Internet infrastructure. See Dan Hunter, "Cyberspace as Place and the Tragedy of the Digital Anticomons" (2003) 91 California Law Review 439, 447–52; Treppoz (n 14) 280.

perspective.[41] Additionally, in the case of Internet crimes, the state's competences are limited to exerting a specific influence on the behaviour of individuals. The point of reference is not cyberspace[42] but rather the social situation—or, more broadly, the crime in which the individual takes part. Cyberspace as such is not subject to regulation—just as the seas are not regulated—but rather only the behaviour of seafarers sailing on ships.[43]

The use of cyberspace for undertaking behaviours, however, results in the transformation of the social meaning of the elements of a social situation that constitutes jurisdictional links. It is possible to indicate here three important modifications.

First, as a rule, in the case of cybercrimes, the physical location of the perpetrator is different from where the socially significant effect of their operation occurs.[44] In the pre-Internet era, the behaviour of the perpetrator affected above all the people around him. Hence, the territory in which he undertook the behaviour was a very important jurisdictional link. Currently, the fact of commission might be invisible to nearby surroundings. A link in the form of the place that a behaviour occurs therefore loses its significance.

Second, the range of the social effects of crime is now widening, becoming available to a wide and undefined number of potential recipients.[45] An example is that placing illegal software on a public Internet server has social effects in many

[41] Such approach is widely represented in the literature. See *inter alia*: Goldsmith (n 26) 32: "Transactions in cyberspace involve real people in one territorial jurisdiction either (i) transacting with real people in other territorial jurisdictions or (ii) engaging in activity in one jurisdiction that causes real-world effects in another territorial jurisdiction. To this extent, activity in cyberspace is functionally identical to transnational activity mediated by other means, such as mail or telephone or smoke signal.". See also: Joel P Trachtman, "Cyberspace, Sovereignty, Jurisdiction, and Modernism" (1998) 5 Indiana Journal of Global Legal Studies 561, 568. In opposition, see: Post (n 30) 1365.

[42] There is also a different approach in the literature, see, among others: Zekos (n 13) 1–2.

[43] The metaphor of the sea is very often used to describe cyberspace, see, among others: Treppoz (n 14) 273–4; Kohl (n 18) 40; Kulesza (n 2) 19.

[44] Michael E O'Neil, "Old Crimes in New Bottles: Sanctioning Cybercrime" (2009) 9 Geo Mason L Rev 237 263.

[45] Zekos (n 13) 6; Mike Keyser, "The Council of Europe Convention on Cybercrime" (2003) 12 J Transnational Law & Policy 287, 294.

countries. This example, in turn, entails a situation in which many states can exert a claim to regulate or impose punishment for the behaviour of the individual.

Third, a characteristic feature of the behaviours undertaken on the Internet is the lack of predictability of the range of consequences and of their multiple locations.

The change in the spatial scope of the social effect of the perpetrator's behaviour and the unpredictability of this scope entail the modification of the scope of state competences.[46] The state claims the right to regulate or punish a given behaviour due to the influence that this behaviour has on social relationships covered by the protection of that state. Although in the pre-Internet era, the physical presence of the perpetrator determined the extent of such influence, physical presence has marginal significance at present.

In the course of the validation and penalisation discourse, traditional arguments (*e.g.*, referring to the place of behaviour) and characteristics for online realities (*e.g.*, based on the location of the server) are intertwined with each other. All of them must be evaluated from the perspective of the interests they describe. Thus, one can formulate reasonable arguments— subject to weighing and considering the interests of both states and individuals—in isolation from rigid jurisdiction rules.

The adoption of the above approach appears justified for two reasons.

First, in the absence of a set hierarchy of jurisdictional principles, entering individual elements of the social situation into their content is pointless. Doing so only leads to the blurring of the differences between important and negligible social interests. At the same time, it does not constitute any added value.

Second, one cannot lose sight of the issue considered here, which concerns the power of the state over the individual. The discursive approach allows for the inclusion of negative arguments based on the values recognised by international law, which are protected by human rights.

The use of a single category (an interest) instead of many jurisdictional principles allows for the construction of universal and more-flexible validation and penalisation arguments. Determining the interests that underlie such arguments enables the meaningful weighing of such arguments. Bearing in mind the above, it appears necessary to transfer the analysis from the level of principles to the level of jurisdictional links (interests).

Further considerations are addressed from the perspective of the two functions that these links perform in the process of introducing and enforcing penal regulations. First, from the perspective of competence to legislate (legislative jurisdiction), they are elements of validation arguments decisive for the effectiveness

[46] See *inter alia* Daskal, Borders (n 18) 185–6.

of norms binding the individual. Second, in terms of the right to enforce the law (*ius puniendi sensu stricto*), they justify extending court jurisdiction to the perpetrator and imposing a penalty on him under the majesty of law. Such distinction is important because in the doctrine of criminal law, the distinction between the validation and penalisation arguments is often omitted. This omission, in turn, results in the transfer of structures based on the fiction of the place committing the act (effective territoriality)—which constitutes an element of penal argumentation—to the plane of validation.

B. Jurisdictional links as validation arguments

A feature of a legal norm is validity.[47] The fact that the standard is in force implies an obligation of the individual to proceed in a certain manner, as described in the standard. The question of whether (and to what extent) a given sentence of a directive character is binding is determined by the validation grounds (arguments).[48] Indication can be made here to establish a norm in accordance with the procedure,[49] but also—most important from the perspective of the considerations discussed here—acting within the limits of the competence of the state.[50] Such limits are defined precisely by referring to jurisdictional links.

To regulate a given social situation, the legislator uses the following (simplified) argumentation.

1. A potential social situation X is characterised by a connection with the state of Y either because the individual performing the behaviour is a citizen of state Y; because behaviour leads to interaction with the citizens of state Y; or because the place of behaviour is the territory of state Y.

2. Hence, the potential social situation of X remains within the competence of the State of Y.

3. Thus, state Y has the right to regulate how the individual shall behave in the potential social situation of X.

To make a reasonable decision about the spatial scope of the norm, it is necessary to construct such a validation argument, which is effective at least when

[47] Kelsen (n 21) 267.
[48] Grabowski (n 39) 445; Zając (n 15) 38.
[49] Grabowski (n 39) 489.
[50] Zając (n 15) 39.

the perpetrator behaves.[51] Because the norm is to provide specific instructions to an individual, its content cannot remain undefined for the moment of action or omission. Enforcement of law against an individual is not meant to build the state's international position but rather to regulate social relationships. Such an approach translates directly into the criteria for the admissibility of validation arguments raised in the discourse. They can only be based on circumstances that are recognisable to a model citizen at the time of his or her behaviour.

Recognition does not mean, however, the existence of such a circumstance at the moment of action or omission. In practice, however, it is occasionally possible to predict future interference that will occur with high certainty. This accuracy will occur in two cases.

First, the perpetrator acting outside the territory of state X makes a targeted attack on objects located in its territory.[52] For example, knowing where the server is located and who owns the data, a hacker might try to break into an Internet server located in the US to destroy important data and cause harm to the property of an American firm.

Second, the act of the perpetrator can, by its very nature, be associated with the induction of a specific consequence, the extraterritorial nature of which will be highly probable. For example, a person can create an online auction in which one can buy material promoting Nazism,[53] and the auction website will be available from anywhere in the world. In both cases, it will be possible to determine at the moment of behaviour that the behaviour of the perpetrator will affect the social situation of another country. In these cases, constructing a valid validation argument is not excluded, because of predictability of circumstances, which constitutes a jurisdictional link.

In contrast, all elements of a social situation whose future occurrence is not predictable for a model citizen cannot form the basis of an effective validation argument. Because the content of the norm is to shape the perpetrator's behaviour,

[51] Carly Henek, "Exercise of Personal Jurisdiction Based on Internet Web Sites" (2000) 15 St John's Journal of Legal Commentary 139, 145; Scott Isaacson, "Finding Something More in Targeted Cyberspace Activities", (2016) 68 Rutgers U L Rev 905, 914.

[52] Cindy Chen, "United States and European Union Approaches to Internet Jurisdiction and Their Impact on E-Commerce" (2004) 25 U Pa J Int'l L 423, 431; Treppoz (n 14) 282; Emily Lanza, "Personal Jurisdiction Based on Internet Conracts" (2000) 24 Suffolk Transnational L Rev 125, 127; *Calder v Jones*, 465 U.S. 783 (1984) <http://cdn.loc.gov/service/ll/usrep/usrep465/usrep465783/usrep465783.pdf> (accessed 5 November 2018); Suomputer Law Review and Technology Journal 2003, vol. VIII at 51; Frank B Arenas, "Cyberspace Jurisdiction and the Implications of Sealand" (2003) 88 Iowa L Rev 1165, 1186; Isaacson (n 51) 919.

[53] *Yahoo! Inc. v La Ligue Contre Le Racisme et L'antisemitisme* 433 F.3d 1199 (9th Cir. 2006) <https://caselaw.findlaw.com/us-9th-circuit/1144098.html> (accessed 5 November 2018) (*Yahoo! v LICRA*). See also: Allen (n 3) 70–5; Hathaway (n 15) 1186; Kohl (n 18) 93.

its validity must be based on such relationships between deed and territory; that is, they must be noticeable *ex ante*.

The catalogue of validation arguments (jurisdictional links) is open. Nothing prevents a state from claiming to regulate a given social situation, for example, because that situation is connected with a high probability of causing some type of effect within its territory.[54] International law does not introduce any limits in this area. Many states might theoretically invoke many different elements of the same social situation and, on this basis, make claims to regulate this situation.

C. Jurisdictional links as penal arguments

The occurrence of jurisdictional links also justifies the imposition of punishment on a person who has committed violations of the applicable law. The state claims the competence to punish specific behaviours because of the negative effect they had on the social relationships under its protection. In this case, the state does not fulfil the right to normalise the social situation but only the competences to perform the functions of criminal law: retributive, compensatory, and protective functions.[55]

In the case of *ius puniendi*, the argumentation for the existence of competences on the side of the state is constructed a little differently:

1. An act of perpetrator X is characterised by a relationship with state Y either because the perpetrator X is a citizen of state Y; because the act violates the interests of state Y or interests protected by state Y; or because the place of committing the act or its consequences is the territory of state Y.
2. Hence, the act of perpetrator X remains within the competence of the State of Y.
3. Thus, state Y has the right to impose punishment on the perpetrator X.

As a social phenomenon, crime is not limited to the behaviour of the perpetrator. It consists of several other circumstances, such as the result or the fact of victimisation, which have a subsequent nature. Although they influence the

[54] Chen (n 52) 435.

[55] Jan Jodłowski, *Zasada prawdy materialnej w postępowaniu karnym. Analiza w perspektywie funkcji prawa karnego* (Wolters Kluwer 2015) 248–312; Christopher W Mullins and Dawn L Rothe, "The Ability of the International Criminal Court to Deter Violations of International Criminal Law: A Theoretical Assessment" (2010) 10 Int'l Crim L Rev 771, 776–84; Peer Stolle, Tobias Singelnstein "On the Aims and Actual Consequences of International Prosecution of Human Rights Crimes" in Tobias Singelnstein *et al* (eds), *International Prosecution of Human Rights Crimes* (Springer 2007) 38–56; George P Fletcher, *Basic Concepts of Criminal Law* (OUP 1998) 12.

assessment of the behaviour of the perpetrator, they are not completely dependent upon him from the perspectives of place and time.

Responding to the need for punishment, the state acts reactively to the behaviour that has already occurred. There is no need to regulate the behaviour of the perpetrator *ex ante* (in fact, doing so is impossible), but only to apply to that behaviour an appropriate means of response, ensuring the implementation of the criminal law function (*ex post* perspective). The social situation lost its potential character and became real. All important elements that can indicate the relationship of the event with the state have occurred and are possible to prove. There is no uncertainty, which was an inseparable feature of a potential social situation.

Creating penalisation arguments based on circumstances that are not recognised by the perpetrator *ex ante* does not necessarily mean excessive restriction of the rights of the individual. The condition for imposing a penalty is to show that the perpetrator act was an infringement of the law binding him at the time of the action or omission.[56] For this purpose, it is necessary to perform validation argumentation to determine the content of the norm regulating a given behaviour. If it is determined that the perpetrator had the right to behave as he behaved, imposition of punishment will be impossible. Individual legal systems even introduce special legal instruments, collectively referred to as the requirement of double criminality. It follows from the above that each state can exercise its competences to a large extent but only on the condition that the basis for its assessment will be a standard coinciding with the norm binding the individual at the time of action.

V. Jurisdictional links included in the jurisdictional principles as validation and penalisation arguments

Analysing the normative content of the jurisdictional principles shaping the scope of the criminal jurisdiction of individual states, it is possible to derive those elements of social situations from them that are in fact jurisdictional links.[57] To switch from a model based on principles to a discursive model, which is based on

[56] Zając (n 15) 260.

[57] See also: Daskal, Borders (n 18) 227.

weighing the significance of links, it is necessary to extract the jurisdictional links from the jurisdictional principles.

A. Territoriality principle

The basis of the principle of territoriality is the space in which an act is committed or, more broadly, a crime occurs. Thus, the territory plays the role of a jurisdictional link. In the course of development, the principle of territoriality has lost its homogeneous character. Currently, its two variants are distinguished: simple and effective.[58]

In the case of the simple variant of the territoriality principle, the territory is understood as a place of a presence of the perpetrator at the time of his act or omission. From this perspective, a person who from his flat in Berlin publishes paedophile content on an American server commits his or her action on the territory of Germany. The link in the form of territory thus understood constitutes an extremely strong basis for validation and penalisation arguments. The state, as a sovereign entity, has as a rule the right to supervise all aspects of social life that occur within its borders.[59] In the pre-Internet era, the effects of the crimes committed above all affected the closest surroundings of the place in which the perpetrator acted. Thus, the place of action or omission coincided with the place in which the public order was breached. These points all speak in favour of granting a territorial connection a very significant meaning. Not without significance is the fact that a territorial connection is also characterised by the highest stability and recognisability.

In addition to the principle of territoriality in a simple approach in the doctrine of international law, the principle of effective territoriality also developed.[60] The jurisdictional link in the form of territory is understood here as the place in which the consequences of a specific act or abandonment of the perpetrator occur. In this case, the states use the fiction of committing an act on their own territory, citing other circumstances than the place of the perpetrator's physical presence.[61]

Classically defined territoriality refers to one circumstance—the presence of the individual in space. Within this approach, there can be no conflict of jurisdictions—science does not address cases of bilocation. The case of effective territoriality is different; the relationship between crime and state is to be

[58] August (n 3) 537; Ryngaert (n 5) 76; Keyser (n 45) 300; Hathaway (n 15) 1840; Kohl (n 18) 24.

[59] *Extraterritorial* (n 4) 446; Treppoz (n 14) 275; Zekos (n 13) 7; Kriangsak Kittichaisaree, *Public International Law of Cyberspace* (Springer 2017) 24.

[60] *Tallin* (n 2) 57; Hafen (n 6) 216; *Extraterritorial* (n 4) 446; Ireland-Piper (n 5) 78; Post (n 30) 1381–4; Podgor (n 11) 730; Kohl (n 18) 89–94.

[61] Ryngaert (n 5) 75.

demonstrated by the occurrence of a specific effect in its territory. Such criteria are extremely diverse. For example, such an occurrence might take the form of, among others, the place of damage,[62] the location of the server on which the data are stored,[63] or the availability of certain content posted on the Internet from the territory of the state.[64]

The diverse nature of the elements that constitute the principle of territoriality in its effective approach will affect the diversification of the strength of the validation and penal arguments based on them (the method of weighing them will be presented in the last part of the work).

As indicated above, the essential feature of the jurisdictional link constituting the basis of the validation argument is its recognisability at the moment of undertaking the behaviour. Such a circumstance is certainly a place to evaluate the behaviour.[65] However, there is a serious doubt concerning whether these elements of the social situation can be treated similarly. This doubt forms the basis for links drawn from the principle of effective territoriality. They are subsequent to behaviour. Here, the important element connecting behaviour with the territory is the change actually caused by the perpetrator's previous behaviour. For obvious reasons, previous behaviour cannot be simultaneous with the behaviour itself. The circumstance that underpins the principle of effective territoriality often cannot be the basis for an effective validation argument.

Circumstances derived from the content of the principle of effective territoriality, however, can successfully co-create penalisation arguments. The essence of a crime is to do evil, which is its consequence. The enforcement of the punishment is subsequent to the act. It only requires demonstrating that the behaviour has had a specific effect on internal social relationships. The space that this evil affects is a very important point of reference from the perspective of criminal law objectives. After all, it is the community of this place that has been harmed, and it is the prerogative of the community to satisfy the sense of justice. The stronger the influence of the crime on a given country's population, the stronger the claim to impose a penalty for it.

B. Personality principle

The basis of the personality principle is the relationship existing between the individual and the state. That relationship has an autonomous character and is independent of the place in which an individual is at the time of action

62 Chen (n 52) 435; Treppoz (n 14) 276.

63 See *inter alia*: Goldsmith (n 26) 21.

64 *Yahoo! v LICRA* (n 53).

65 Kohl (n 18) 144.

or omission. The personal bond is considered here from the perspective of the relationship between the addressee of the norm and the state.[66] The person against whom the state claims the right to regulate behaviour or impose punishment must have a feature that testifies to its relationship with the state exercising competences. In fact, every state claims a competence to exercise to some extent power over individuals who have a special relationship with them. One aspect of this approach is the extension of state authority to individuals' behaviour abroad.[67] States thus discipline their citizens or other dependent individuals.

In the case of the analysed link, there is a noticeable difference between its significance as an element of the validation argument and the significance of a penalising argument based on it.

A personal relationship can become the basis for a validation argument. Its existence is recognisable *ex ante*—the person executing the behaviour will usually be aware of his nationality or permanent residence. However, this argument is relatively weak. In most cases, in the course of a validation discourse, it will give way to an argument based on a territory connector. It is impossible to imagine that a person travelling from State X to the State of Y would somehow transfer the entire legal system binding that person in his homeland. One cannot require the person to, for example, drive the vehicle with the left hand in right-hand traffic (this point is important, for example, in determining who is the perpetrator of a car accident). Behaviour must be consistent with how the community operates.

In practice, however, there will be a certain group of norms, the extraterritorial effectiveness of which will be possibly based on the personal link. They refer to situations in which the behaviour of the individual is morally reprehensible from the perspective of the value system of the forum state and at the same time constitutes an expression of an individual decision of the perpetrator (and is not the result of necessary involvement due to, for example, the nature of social life). Such situations include, for example, acts of a paedophile nature, from which the state will be able to ban its citizens.[68]

The significance of the penal argument based on personal bond is different. The argumentation for extending the law is additionally strengthened by various circumstances. The community has a strong need to account for the evil performed

[66] *Tallin* (n 2) 62.

[67] Lanza (n 52) 126; Ryngaert (n 5) 90; Hafen (n 6) 218; Titi Nguyen, "A Survey of Personal Jurisdiction Based on Internet Activity: A Return to Tradition", (2004) 19 Berkeley Tech LJ 519, 520; Keyser (n 45) 315; According to the Article 22 of the Convention on Cybercrime. Reference, ETS No.185. Opening of the treaty, Budapest, 23/11/2001: "Each Party shall adopt such legislative and other measures as may be necessary to establish jurisdiction over any offence established in accordance with Articles 2 through 11 of this Convention, when the offence is committed: [...] by one of its nationals, if the offence is punishable under criminal law where it was committed or if the offence is committed outside the territorial jurisdiction of any State".

[68] Keyser (n 45) 306.

by a person who is part of such a community. Moreover, in view of the growing respect of the principle of *ne bis in idem*, associated with the ban on the issue of own citizens,[69] imposing a punishment is often the only instrument to satisfy the social sense of justice.

C. Passive personality principle

The principle of passive personality is in a sense the reverse of the principle of active personality. In this case, the competence for regulating and imposing punishment is based on the relationship between the State and the entity whose good has been violated by the offense (victim). The point of reference is here infringed legally protected values (life, health, property, but also the right to be forgotten),[70] which can be linked to a specific state by a person who is a disposer of such values (*e.g.*, the owner of a thing).[71] The state claims the right to protect the interests of entities that undertake their activities under the protection of such state's jurisdiction. The scope of this protection is defined not only by the place of specific activity or passivity but also by the nationality of the persona or the place of registration of the legal entity.[72]

The use of the abovementioned personal link as the basis for the validation argument entails significant problems. The perpetrator until the moment of action does not remain in any relationship with the forum state. At the same time, it is often unpredictable ex ante whose property will be involved in a given future social situation. In many cases, the perpetrator has no real opportunity to recognise the nationality of the person with whom he or she engages in a specific social situation, *e.g.*, proposing the presentation of pornographic content. This ambiguity is particularly common in the case of interaction in cyberspace, in which the principle is anonymity. Moreover, information distributed via the Internet influences an undefined group of people, each of whom can be subject to the protection of a different legal system. Even when it is possible to unambiguously determine ex ante the circle of entities to which a given information arrives, the question arises concerning the possibility of a sensible solution towards which the state will have the competence to regulate such behaviour.

The doubts outlined above are missing if the analysed link is considered the basis of the penal argument. The claim to impose a penalty will be justified by the

[69] Ryngaert (n 5) 90; *Extraterritorial* (n 4) 448.

[70] Daskal, Borders (n 18), 209.

[71] Watson (n 6) 18; Hafen (n 6) 218; McCarthy (n 6) 301; August (n 3) 541; *Tallin* (n 2) 64.

[72] *Extraterritorial* (n 4) 451; Zając (n 15) 423.

fact that the entity being under state protection is harmed.[73] However, such claim is no longer as strong as in the case of effective territoriality. The crime committed there indirectly affected the national public order. In the case of a connector based on the "origin" of a legal good, such a circumstance does not constitute a connecting entity.

D. Protective principle

The protective principle is based on a jurisdictional link in the form of essential state interests.[74] The criterion of the disposer of legally protected value, inscribed in the content of the principle of passive personality, is here supplemented with additional elements clarifying the character of the good itself. In the classical approach, the analysed principle referred to those interests whose security conditions the existence of the state as a sovereign political organism.[75] In the second half of the twentieth century, individual countries also began to refer to the need to protect values of a supra-individual character, whose security lies in the interest of the state due to the benefits (for example, securing the market balance by prohibiting the corruption of own entrepreneurs).[76]

Depending upon the type of interest appearing in a given social situation as the base of the argument, the strength of the validation or penalising argumentation is different.

The state has strong competences to normalise these behaviours, which are directed at interactions with legally recognised interests of such a state.[77] There is no doubt that state X is entitled to regulate the question of addressing data on government servers or to determine the procedure for obtaining access to such data. Only state X is interested in the proper protection of this sphere of public life. At the same time, there will be no element of uncertainty here in principle.

[73] Regula Echle, "The Passive Personality Principle and the General Principle of Ne Bis In Idem" (2013) 9 Utrecht L Rev 56, 57.

[74] *Extraterritorial* (n 4) 451; Matthew Garrod, "The Protective Principle of Jurisdiction over War Crimes and the Hollow Concept of Universality" (2012) 12 International Criminal Law Review 763, 776; Edward T. Meyer, "Drug Smuggling and the Protective Principle: A Journey Into Uncharted Waters" (1979) 39 Louisiana L Rev 1189, 1190; Slomanson (n 5) 252; *Tallin* (n 2) 63.

[75] Hafen (n 6) 217.

[76] *Extraterritorial* (n 4) 454. The global nature of the interests (legal goods) determined in this way means that in the analyzed scope the protection principle is sometimes identified with the principle of universal jurisdiction. However, this approach does not seem legitimate. In the case of a protective principle, the state relies on a supranational good and undertakes to establish specific regulations in this respect. In the case of a universal rule, these regulations are derived from customs and general principles of law.

[77] *Extraterritorial* (n 4) 451; Ryngaert (n 5) 97; Jared Beim, "Enforcing a Prohibition on International Espionage" (2018) 18 Chi J Int'l L 647, 670.

State legal goods have a special character and thus are recognisable—for example, due to the domain address or by the introduction of used signs (for example, by the official stamp "secret").

The situation will be slightly different in the case of supra-individual goods, somehow inscribed in the content of the existing jurisdictional principle. They are not reflected in objectively perceived reality. A legal good such as "market equilibrium" lacks a tangible substrate, which can lead to a lack of identifiability of the relevant link. Moreover, based on the jurisdictional link thus described, many states can claim the competence for regulation. In fact, no state is entitled to that competence more than another, because all states indicate the same jurisdictional link.

The doubts outlined above are weakened if the link-based interest protected by the state is considered an element of the penal argument. In both the case of state interest and those cases of a transnational character, the state has a strong non-exclusive claim to impose a penalty. In the first case, that claim will result from the need to pay for the attack on the political organism. In the second country, the forum will appear on behalf of the international community interested in securing a particular interest.

E. Vicarious jurisdiction

The essence of the principle of substitute criminalisation is to impose a punishment in the name of the originally authorised state.[78] Under this principle, the forum state imposes a penalty if the perpetrator of a crime committed abroad is currently on its territory.[79] The circumstance that is the basis of the competence, remain in the country—it is not related to the perpetrator's behaviour itself, or even with its result. Therefore, the circumstance cannot be the basis for constructing an effective validation argument.

A penal argument in fact refers to the element of territory.[80] In this case, the state grants itself the power to punish the perpetrator due to the need to protect its own society and public order against further attacks by the perpetrator remaining

[78] Ryngaert (n 5) 102–3; Extraterritorial (n 4) 452.

[79] Ryngaert (n 5) 106.

[80] Michał Płachta, "Zastępcza represja karna w prawie polskim" in Piotr Hofman ski and Kazimierz Zgryzek (eds), *Współczesne problemy procesu karnego i wymiaru sprawiedliwości: Księga ku czci Profesora Kazimierza Marszała* (Wydawnictwo Uniwersytetu S la skiego 2003) 353.

within its borders. By imposing a punishment, the forum state also avoids the opinion of a rogue state that provides protection to criminals.[81]

F. Universal jurisdiction

The tragic events of the twentieth century brought a change in the perception of the role of states as guardians of value. Although the state was primarily responsible for safeguarding interests in its "own backyard", wars and humanitarian tragedies somehow forced states to change this approach.[82] States have taken responsibility for protecting certain universally recognised values (basic moral values) against the most serious categories of violations.[83] Thus, recognition has acquired a new jurisdictional link that can be defined as "the universal interest of the international community".[84]

Considering this link, the basis for the validation and penalisation argument, unlike other jurisdictional linkers, one can impute the character of prohibited or prescribed behaviour into the jurisdictional link itself. Each state has the right to punish committing genocide, for example.[85] However, no state has the power to shape universal standards of conduct in this respect; this point is described in international law.[86] In this case, the role of the state is basically only in the insertion of a specific norm of international law into national legal orders. The analysed link is therefore the basis for the penalisation argument.

VI. New jurisdictional links as validation and penalisation arguments

Observing the problem of crime committed with the use of the Internet, individual states began to claim the right to regulate and punish behaviour undertaken in cyberspace based on modified jurisdictional rules. This approach essentially meant extending the application of the norms of a given country by

[81] Beim (n 77) 660.

[82] Mark Lewis, *The Birth of the New Justice: The Internationalization of Crime and Punishment 1919-1950* (OUP 2014) 80.

[83] Ryngaert (n 5) 114; Ostropolski (n 34) 47–8.

[84] Ryngaert (n 5) 106, 114; The Princeton (n 10) 28; Hafen (n 6) 219; Bassiouni (n 4) 280.

[85] Ryngaert (n 5) 116; Cynthia Sinatra, "The International Criminal Tribunal for the Former Yugoslavia and the Application of Genocide" (2005) 5 Int'l Crim L Rev 417, 417–8.

[86] August (n 3) 542.

reference to circumstances such as the location of the server or the language of the message being transmitted.[87]

The considerations presented here discuss the possibility of treating some of the "new circumstances" as autonomous jurisdictional links.[88] However, they are not comprehensive. The purpose of the argument is to show that the circumstances can be described using dogmatic tools and subjected to rational weighing.

A. Jurisdictional links related to computer infrastructure

Some states claim the right to regulate those activities performed in cyberspace that use elements of the IT infrastructure located on their territory. The claim is primarily indicated in this section for servers that collect data[89] (*e.g.*, illegal duplicate copies of software) and transmission devices.[90]

The presence of devices used to commit an act is used to prove the relationship of behaviour with the territory. This connection in turn allows us to construct effective validation and penalisation arguments, which refer to the interference of behaviour in the internal relationships of the state. The problem is, however, that the significance of the social situation in cyberspace focuses on the sender and recipient of information. The mere "storage" of a digital record, often encrypted, does not interfere with social relationships that are protected by the state of the server's location.[91] Not without significance is the fact that specific data can be stored simultaneously on many servers, located in different countries.[92]

The possibility of using the location of ICT infrastructure elements as a validation argument appears significantly limited due to the unpredictability of the transmission route or the storage location. The average user of the network will not even be able to determine the path of data transfer or where they are physically stored.[93]

The possibility of constructing penal arguments is slightly different. There will be no obstacles related to the lack of predictability here. Both the location of

87 Daskal, The Un-Territoriality (n 18) 355.

88 In the course of deliberations, the technical aspects of the Internet were omitted. In this respect, see, among others: Chen (n 52) 426–7.

89 *Yahoo! v LICRA* (n 53). See also: Chen (n 52) 447; Goldsmith (n 26) 21; Jennifer Daskal, "Law enforcement Access to Data Across the Borders: The Evolving Security and Rights Issues" (2016) 8 J National Security Law & Policy 473, 490.

90 Hunter (n 40) 477; Zekos (n 13) 14; Daskal, Borders (n 18) 188; Daskal, The Un-Territoriality (n 18) 326; O'Neil, (n 44) 254.

91 Daskal, The Un-Territoriality (n 18) 371.

92 Kirstern E. Eichensehr, "Data Extraterritoriality", (2016) 95 Tex. L. Rev. 145, 145; Daskal, Borders (n 18) 223.

93 Burk, (n 2) 123; Zekos (n 13) 15; Daskal, The Un-Territoriality (n 18) 366; Grootes (n 12) 17: "There really is no "there" there in cyberspace, but mere geographic probabilities or uncertainty".

servers and the means of data transmission can be proved during the process.[94] Moreover, the state in which the server is located has the potential to effectively block illegal content—for example, in criminal proceedings.[95]

B. Place of domain registration

Some states try to extend their jurisdiction to domain registration criteria.[96] An example is the USA.[97] Each webpage is referred to as an address ending in a so-called the top-level domain (TLD).[98] Such domains are divided into domestic domains (for example, .pl and .de) and functional domains (for example, .com and .org). The management of the TLD is performed by the ICANN organisation.[99] It has been established and operates in the US, giving the US real control over cyberspace.[100] In addition, individual countries manage national domains to a certain extent by controlling national DNS servers.[101]

Considering the criterion of the place of domain registration an element of the validation argument, wherever the address of the website is connected to the functional domain, determining the place of domain registration is virtually impossible or very difficult.[102] Lack of recognisability in this case excludes the possibility of constructing validation arguments. When the website address is based on the national domain, *prima facie* it is possible to predict the country of its registration.[103] In each case, however, the circumstance creates a relatively small relationship between the social situation and the state. A given website operating based on a specific domain lacks any social significance. Behaviour that occurs with the use of a specific domain does not interfere in principle with the internal relationships of the state of its registration.

Concerning penalisation arguments, nothing prevents such arguments from being constructed based on both the domestic and functional domains. However, the question concerning the power of such arguments remains valid. This claim

[94] Xinbao, Ke (n 22) 43.
[95] Zekos (n 13) 4.
[96] Thomas R Lee, "In Rem Jurisdiction ind Cyberspace" (2000) 75 Wash L Rev 97, 116; Hathaway (n 15) 1827.
[97] Edward Lee (n 18) 1332.
[98] Thomas R Lee (n 96) 102.
[99] Hunter (n 40) 479; Edward Lee (n 18) 1332; Xinbao, Ke (n 22) 43.
[100] Xinbao, Ke (n 22) 51; Kulesza (n 2) 87.
[101] Edward Lee (n 18) 1332; Thomas R Lee (n 96) 100.
[102] Chen (n 52) 432; Hathaway (n 15) 1828.
[103] Grootes (n 12) 17;

will be strengthened if it is shown that the state in question has a real possibility of blocking the domain.

C. Accessibility of data

The Internet provides one the possibility of unlimited information (data) distribution. However, some communications might be undesirable and fuel dangerous social phenomena (*e.g.*, glorifying Nazism or insulting religious feelings). Therefore, it is claimed that the state has the right to normalise and punish behaviour that leads to the creation of a message available in its territory.[104]

There are no obstacles preventing reference to the criterion of the availability of a message as a validation or penalisation argument.[105] It should be assumed that a message posted on the Internet reaches every country.[106]

The inability to reasonably rely on the accessibility argument stems not so much from the lack of predictability as from the inconclusive character. Thus, a claim to normalise or punish the same behaviour can be reported by several dozen countries. In this situation, it is impossible to determine who has a stronger power.[107]

In practice, reference to the criterion of the availability of a message will be reduced to demanding such a narrowing of its scope that it will not be available in the territory of a given country.[108] Such a request is, however, made not under criminal law but under private or administrative law. Moreover, the recipient of the request will not be the perpetrator here but rather the person who has the

[104] *Dow Jones and Company Inc v Gutnick* [2002] HCA 56; 210 CLR 575; *Yahoo! v. LICRA* (n 53); ECJ Hedjuk C-441/13 (2015). See also: Allen (n 3) 74; Chen (n 52) 434; Treppoz (n 14) 277; Edward Lee (n 18) 1288; Lanza (n 52) 132; Kohl (n 18) 47.

[105] The criterion for the availability of the message was used by France in the dispute with Yahoo! (*Yahoo! v LICRA*, n 53). The portal offered Nazi souvenirs in its auction house, the sale of which was prohibited in France. Among the many steps to block the auction, the French prosecutor's office also initiated criminal proceedings against Timothy Koogle, a former executive director. See, among others: Chen (n 52) 448; Grootes (n 12) 3–4; Kohl (n 18) 100.

[106] *Dow Jones v Gutnick* (n 104); Chen (n 52) 423; Zekos (n 13) 6; Kulesza (n 2) 15.

[107] For these reasons, the accessibility criterion is questioned as an enabling condition of regulation. See ECJ L'Oreal C-324/09 (2011). Studies suggest replacing them with the criterion of purpose. See among others Treppoz (n 14) 282; Goldsmith (n 26) 15; Isaacson (n 51) 922.

[108] *Yahoo! v LICRA* (n 53). See also: Chen (n 52) 447; Treppoz (n 14) 277–8; Daskal, Borders (n 18) 195; Allen (n 3) 83; Charmaine H Perdon, "The Regulation of Cyberspace" (1999) 73 Phil LJ 569, 595–7; Kohl (n 18) 228–9.

technical means of modifying such a coverage (*e.g.*, administrator of the auction site on which souvenirs from the period of Nazi Germany are sold).[109]

D. Language of the message

Due to the undecidable nature of the jurisdictional link in the form of a message, some countries have begun using the criterion of language.[110] Language can be a rational criterion that narrows the circle of information recipients and clarifies the scope of its effect. Moreover, it often coincides with the national criterion.

There are no formal obstacles preventing the construction of a meaningful language-based argument—both in the validation discourse and penalisation. The question is about their persuasive power. The language of the message should be considered in those cases in which it clearly identifies the information recipient. The situation is complicated in the case of common languages, in particular English, which is currently the *lingua franca* of cyberspace. Against this background, a problem of languages is also recognised as official in a larger number of countries (*e.g.*, Spanish, English, and German). Each of them will be able to rely on the same circumstance as a validation or penalisation argument, which will create further conflicts of jurisdiction.

VII. Process of weighing arguments based on a jurisdictional link

In international law, no universally accepted hierarchy of jurisdictional links has been developed. It appears necessary to present the method of weighing their significance. It will thus be possible to compare the strengths of the arguments that are based on them. The weighing method will be slightly different in the case of both types of arguments.

In the case of validation arguments, the weighing process must lead to an unambiguous determination of which directive of conduct applies to the person in a given classification. The effect of the validation discourse should be the answer to the question, "What is the norm applicable in situation X?" The person undertaking the activity must be able to determine the unambiguous content of the law binding him or her. This requirement does not mean that there will be any

[109] Chen (n 52) 452; Edward Lee (n 18) 1330.

[110] Treppoz (n 14) 282; Susan N Exon, "Personal Jurisdiction: Lost in Cyberspace?" (2003) 8 Computer L Rev & Tech J 21, 34.

certainty here. Rather, the need to unambiguously determine the content of one binding standard will be associated with a high risk of error.

Which of the arguments (interests) in favour of the right to regulate a given situation by state X will be considered the most important must be set *in concerto*. This process occurs in two stages.

First, the nature of the interests must be considered. International law has not developed a precise method of weighing them. There are no categories that can be characterised by greater recognition or recognisability. Moreover, the claim can be supported by many validation arguments, based on various links, which will be followed by their stronger binding power. When assessing the relevance of a country's interest, it is necessary to use the so-called rule of reason,[111] which is derived from four principles of international law: the principle of non-intervention,[112] the principle of equity,[113] the principle of proportionality,[114] and the prohibition of abuse of law.[115] From this perspective, an attempt to normalise a given social situation by the state with a less significant interest (weaker argument) can be treated as a violation of international law (rule of reason) and be the subject of proceedings before the International Court of Justice.

Second, one should verify whether the claim of the state who holds the strongest arguments is not blocked by negative premise. Such a negative premise is the lack of objective predictability (foreseeability) of the occurrence of a given jurisdictional link. If a given circumstance is unrecognisable *ex ante* (at the moment of behaviour), then it cannot be the basis for a validation argument. Consequently, this argument must be rejected, which can lead to a denial of the right to regulate a given behaviour by a state. An attempt to enforce such a denied right might also constitute a violation of international law or human rights.[116]

Note also that for some norms, conflict-of-law rules apply. This point will be true primarily in the case of private law.[117] In this respect, due to the existence of precise collision solutions, reference to the general rule of reason is not necessary.

The process of weighing penal arguments will be slightly different. The effect of penal discourse should be the answer to the question, "can the court of state X punish the perpetrator of the act?" There is no need to explicitly determine

[111] August (n 3) 535; Płachta (n 34) 122; Ryngaert (n 5) 143.
[112] Ryngaert (n 5) 144.
[113] Subir K Chattopadhyay, "Equity in International Law: Its Growth and Development" (1975) 5 Ga J Int'l & Comp L 381, 392; Władysław Czapliński and Anna Wyrozumska, *Prawo międzynarodowe publiczne. Zagadnienia systemowe* (C.H. Beck 2014) 138–42.
[114] Ryngaert (n 5) 148.
[115] Ryngaert (n 5) 150–1.
[116] Zając (n 15) 255–7.
[117] Lanza (n 52) 142; Daskal, Borders (n 18), 182–3.

which state has the exclusive right to punish. International law does not know the universally binding prohibition *ne bis in idem*.[118] Theoretically, it will be possible to punish the perpetrator repeatedly for the same act by various authorised states.

Imposing a punishment on an individual in principle does not interfere in the internal relationships of a foreign country. Against this background, however, there can be a violation of human rights. It is forbidden to punish a behaviour that did not constitute a crime based on the law binding the perpetrator at the time of commitment of the act. The law that binds the perpetrator at the time and place of the act is determined based on the validation arguments. If the punishment is inflicted by a foreign state but for an act that constitutes a crime also based on the law applicable to the perpetrator, there is no breach of the principle of *nullum crimen*.

VIII. Summary

The considerations presented above allow drawing the following final conclusions. First, shifting the burden of validation discourse from the plane of jurisdiction principles to the level of weighing the interests of individual entities allows more efficient settlement of conflicts of jurisdiction. Individual interests are subject to weighing—it is possible to compare their significance. Second, the application of the method based on the weighing of interests allows avoiding the deadlock that occurs when several states invoke the same jurisdictional principle—for example, in the case of effective territoriality. Third, by referring to the rule of reason derived from the four principles of international law recognised by civilised nations, individual states can question the possibility of applying foreign law to the regulation of specific social situations. Disputes can be resolved by diplomatic channels or before international law courts.

[118] Sakowicz (n 8) 89–133; Lech Gardocki, *Zagadnienia internacjonalizacji odpowiedzialności karnej za przestępstwa popełnione za granicą* (Wydawnictwo Uniwersytetu Warszawskiego 1979) 50; Barbara Nita, "Zasada ne bis in idem w międzynarodowym obrocie karnym" (2005) 709 Państwo i Prawo 18; Janusz Tylman, "Środki przymusu w postępowaniu karnym przeciwko cudzoziemcom" in Andrzej Szwarc (ed), *Przestępczość przygraniczna. Postępowanie karne przeciwko cudzoziemcom w Polsce* (Wydawnictwo Poznańskie 2000) 43.

Cambridge Law Review (2019) Vol IV, Issue ii, 29–67

Regulating Virtual Currency Payment Systems

Sijuade Animashaun*

Abstract

This paper examines the functioning of virtual currencies as payment systems through crypto-currency exchanges and the likely impact their integration with traditional payment systems may have on the interdependent global payment systems. Being a potential global transformational phenomenon, should virtual payment systems be regulated like other traditional intermediaries to manage the risks from their operations? Which regulator has the requisite regulatory architecture to comprehend the fast-evolving dynamics of the innovative payment solution and better manage the risks? These are some of the questions attempted in this paper. The paper also examines the role played by central banks as the major regulator of payment intermediaries and their limitations on multinational financial institutions and payment activities. Finally, the paper suggests the adoption of international regulatory bodies as the major regulatory authority for the virtual exchanges in ensuring global cooperation and coordinated implementation of any developed action plan while fostering financial innovation.

I. Introduction

The participation of new financial technology providers in payment and settlement systems is driving a phenomenal change in the financial architecture of global economies.[1] Virtual Currency (VC) Schemes—through their Distributed

* LLB (LASU), BL, LLM (University of Glasgow). Research Assistant at the University of Glasgow. I am grateful to Dr Javier Solana for his supervision and constructive criticisms on the earlier drafts.

[1] See Jun Liu, Robert J Kauffman, and Dan Ma, "Competition, Cooperation, and Regulation: Understanding the Evolution of the Mobile Payments Technology Ecosystem" (2015) 14(5) Electronic Commerce Research and Applications 372–91.

Ledger Technology Blockchain—is the most recent of these innovations and notorious for benefits of transactional speed, efficiency, financial inclusion and most importantly, financial independence because it operates over a peer-to-peer (P2P) network without need for conventional financial intermediaries.[2] VC's new intermediaries—Crypto Currency Exchanges—provide online gateways and storage for VCs and facilitate VC transfers and convertibility with state-issued currencies within the virtual community. These exchanges are evolving rapidly and exploiting the economies of scale and scope of transnational payments processing because they operate on a global scale.[3]

Crypto-currency exchanges, like VC schemes, offer several potential benefits including faster and more efficient transnational payments and settlements at significantly lower costs. These benefits may soon transform the way global payment systems works if VC exchanges become widely used or their services adopted by Traditional financial institutions (TFIs) and other payment service providers (PSPs) engaged in large value cross-border payments and funds transfer.

But these new intermediaries are not flawless, their operations—predominantly outside the regulatory perimeter of traditional banking system—make them vulnerable to the risks of being used as vehicles for financial crimes, particularly fraud through cyber-attacks.[4] Collaborations between multinational TFIs within the banking system and technology providers is popular in both retail and wholesale cross-border payments. The principal drivers of these integrations are majorly the growth of e-commerce and the desire by TFIs to provide faster payment services to a broader demography of clients while reducing operational costs.[5] If (and when) this trend extends to VC crypto-currency exchanges, it could significantly impact the global payment systems through increased efficiency and cross-pollination of risks.

Attempts to regulate the VC intermediaries by national economies have not been unified and arguably inefficient due to territoriality and fragmentation

[2] Volker Brühl, "Virtual Currencies, Distributed Ledgers and the Future of Financial Services" (2017) 52(6) Intereconomics 370.

[3] Dan Awrey and Kristin van Zwieten, "The Shadow Payment System" (2018) 43(4) Journal of Corporation Law 775–816.

[4] International Monetary Fund, *Virtual Currencies and Beyond: Initial Considerations* (January 2016) <https://www.imf.org/external/pubs/ft/sdn/2016/sdn1603.pdf> (accessed 9 August 2018).

[5] Bank for International Settlements (Committee on Payments and Market Infrastructures and Markets Committee), *Fast Payments – Enhancing the Speed and Availability of Retail Payments* (November 2016) <https://www.bis.org/cpmi/publ/d154.pdf> (accessed 9 August 2018).

among national regulators.[6] This is primarily because VCs and crypto-currency exchanges operate on a global scale through the internet and can therefore not be controlled exclusively by a domestic regulator. Concerns about navigation of risks between connected financial systems become particularly amplified by the possibility of a consolidation between TFIs and cryptocurrency exchanges. This is because TFIs are key stakeholders indirectly connecting financial systems in the tightly interdependent global payments network. Any failure from such consolidation may potentially disrupt global financial stability through the spread of systemic failures among the interdependent systems.

For this reason, it's incumbent to consider how best to manage the exposures in the fast-evolving innovative payment intermediary without crippling its potential benefits to the global economy. And more importantly, being a global concern, which regulator possesses the regulatory infrastructure to manage the risks these new entrants pose to the increasingly interdependent global payment system without stifling financial innovation.

A. Aims and objectives

The aim of the paper is to critically analyse the possible impact of VC-intermediaries direct integration with key stakeholders (TFIs) in the global payment systems. The paper also examines how interdependencies among financial systems—driven by financial consolidation-creates an exposure to systemic risks through the indirect interconnectedness of payment systems with same multinational TFIs. The regulators of payment systems are examined with aim of analysing their efficiency in the management of cross-border risks by TFIs. Lastly, the objective of the paper is to argue for the regulation of VC-intermediaries by international bodies in collaboration with domestic regulators in managing systemic risks-particularly fraud from cyber-attacks— that may threaten the potential benefits of the integration between VC-intermediaries and TFIs.

To achieve this aim, the paper will examine the drivers of interdependencies in the global payment systems-particularly financial integration, and the possible benefits and risks of adopting the VC intermediaries as payment solutions providers in cross-border payment and settlements. The cross-pollination of risks between financial systems that may arise from VC-intermediaries consolidation with TFIs

[6] SJ Hughes and ST Middlebrook, "Advancing a Framework for Regulating Cryptocurrency Payments Intermediaries" (2015) 32(2) Yale Journal on Regulation 505.

are examined and a practical approach to managing these potential disruptions is suggested using the instrumentality of international bodies.

B. Value of research

Scholarly discourses on VCs are primarily on its functioning as an alternative payment or financial instrument. While this is laudable, it avoids the potentials of the innovative payment and settlement solution-Blockchain-used by the VC intermediaries in the provision of fast and cost effective cross-border payment services. Two major factors make the consideration of the potentials important presently: (a) there is a growing initiative among large TFIs on possible adoption of VC-Blockchain as an innovative payment and settlement solution to reap the benefits of economies of scale in cross-border payments processing; and (b) several central banks are presently considering the creation of VCs as an alternative payment instrument.

These factors make the financial consolidation between TFIs and Crypto-currency exchanges more likely than ever before if it becomes popular. The attendant risks of such integration are therefore worthy of consideration and more importantly, there is need to examine the regulatory infrastructure presently available to determine which of the regulators have the requisite infrastructure and expertise to contain any potential disruption to global financial stability that might result from this paradigm shift when it occurs. This paper will provide brief insights into the benefits and exposures of such financial integration and serve as a point of reference when considering the appropriate regulator for the financial innovation.

C. Scope of research

Due to the constantly evolving nature of payment systems, there is no comprehensive literature dealing exhaustively with the dynamics of the recent facts considered in this paper. In addition, the relative novelty of VCs and their intermediaries implies that the academic writings on the subject are still growing with not trusted empirical evidence on the phenomenon. The coverage of the paper is to critically analyse the dynamics of VCs as payment systems, not in themselves, but through crypto-currency exchanges. The benefits and risks of financial integration within an interdependent global payment system are also considered through an examination of the drivers that have transformed the global payment infrastructure. Finally, it considers (rather collectively), the functioning of

international bodies and argues for their adoption as the major regulators of VC-intermediaries.

The paper however has limitations, it does not consider the functioning of VCs as payment instruments or financial instruments generally. It also does not cover how VCs or other payment systems are regulated but instead focuses on the principal regulators. Other limitations not mentioned here are stated in the paper.

D. Methodology

The paper adopts the critical analytical approach. It considers the laws, rules, principles and guidelines of payment systems generally. The purpose is to criticise the divergent approaches by domestic regulators to the potentials of VC-intermediaries as a disruptor of global payment systems due to its present limited use. The paper also examines the benefits that VC system's inclusion might sequel for global payment market participants and while acknowledging the possible exposures. Due to the novelty of the research topic, sources researched include books, articles in law journals and other economic journals, publications by regional and international regulatory bodies and other publications in legal and related fields which provide the breadth and depth required for such analytical academic work.

E. Research structure

The research is divided into five sections. This section, Section I, has given a general introduction of the research topic by elucidating the aims and objective, reasons and value of the academic work, methodology and finally, the scope and limitations of the paper.

Section II has three main parts. Part A examines payment systems generally with focus on its core functions in the financial systems. In Part B, this paper examines VCs as a payment system and how it works. Finally, Part C examines whether VCs perform core functions of payment systems. This is achieved through a focus on the functions of crypto-currency exchanges relatable to TFIs.

In Section III, this paper examines how VC intermediaries may impact the global payment systems. Part A examines how global payment systems work with critical attention to the interdependencies of global payment systems. In Part B, the impact of a possible financial consolidation between a crypto-currency exchange and a TFI on the interconnected global systems was considered-particularly

systemic risks. Efficiency and cross-pollination of risks between the financial institutions and indirectly among their financial systems was also examined.

In Section IV, inferences were drawn from the global nature of any impact a consolidation between the VC-intermediary and TFI may have on the global financial systems. The aim of this is to determine who should regulate the VC intermediaries.

Part A examines concisely, whether VC systems require regulatory intervention from the regulators of traditional payment intermediaries. In Part B, focus was on Central banks as the major regulator of payment systems. The limitations to their scope and authority particularly to transnational payments and multinational financial institutions was examined. Finally, in Part C, an argument was made for the adoption of international regulatory bodies as appropriate regulator for the VC intermediaries because of their global operations. In arguing this, focus was on the need for international coordination and cooperation among key stakeholders in the global payment systems in jointly managing global risks. In the end, it was suggested that a joint effort of both national regulators and international bodies was critical to management of the exposures from the innovative payment service providers. I argue that this will guarantee financial innovation that divergent approaches of domestic regulators will otherwise cripple through overregulation.

This paper concludes in Section V by reflecting on the facts considered throughout the paper.

II. Virtual currencies and payment systems

The aim of this Section is to examine whether VCs-through Crypto-currency exchanges perform similar functions as conventional payment system service providers (TFIs). In achieving this goal, the Section is divided into three parts. Part A will focus on definitions of payment systems, its subcategories and core functions of TFIs within the payment network. In Part B, the Section examines VCs as a payment system generally and how it works. Thereafter, Part C considers the question of whether VCs perform similar core functions as payment systems. In this part, the functioning of crypto-currency exchanges is employed to answer the research question.

A. Payment systems

1. What are Payment Systems?

Money plays a crucial role in the overall stability of any economy. It facilitates trade and ensures the smooth running of government and the financial system.[7] Several attempts to define payment systems in legislative instruments and scholarly discourses focus on the banking system which performs core financial functions-including payments processing.

The UK Banking Act 2009 defines payment systems within the banking structure as "an arrangement designed to facilitate or control the transfer of money between banks (and building societies) who participate in the arrangement".[8] The Financial Services (Banking Reform) Act also considers it "a system which is operated by one or more persons in the course of business for the purpose of enabling persons to make transfers of funds, and includes a system which is designed to facilitate the transfer of funds using another payment system".[9] In the same vein, the Committee on Payments and Settlement Systems (CPSS) defined it as "a set of instruments, banking procedures and, typically, interbank funds transfer systems that ensure the circulation of money".[10] In scholarly discourses, they are broadly defined "as a collection of institutional arrangements that facilitate the transfer of funds and other assets in satisfaction of financial obligations".[11]

The most explicit of the definitions is however contained in the Payment Services Directive (EU) 2015/2366 which defines it in Article 4(7) as "a funds

7 Jürgen G Backhaus, "Money and its Economic and Social Functions: Simmel and European Monetary Integration" (1999) 58(4) The American Journal of Economics and Sociology 1075.

8 Banking Act 2009, section 182(1).

9 Financial Services (Banking Reform) Act 2013, section 41.

10 Committee on Payments and Settlement Systems, *A glossary of terms used in payments and settlement systems* (March 2003) <https://www.bis.org/cpmi/glossary_030301.pdf> (accessed 10 August 2018).

11 John Armour, *Principles of Financial Regulation* (1st Ed, OUP 2016) ch 18, citing Andrew Haldane, Stephen Millard, and Victoria Saporta (eds), *The Future of Payment Systems* (Routledge: Abingdon 2007) 2.

transfer system with formal and standardised arrangements and common rules for the processing, clearing and/or settlement of payment transactions".[12]

The nature of payment systems, discernible from these definitions may be summarised as the arrangement for facilitation of funds (in lieu of cash)[13] between a payor and payee for the settlement of financial obligations.[14]

2. Categories of Payment Systems

Save low-value transactions which are majorly facilitated by other payment instruments-credit and debit cards, e-money (including virtual currencies), electronic funds transfer, majority of funds transfer are facilitated through the banking system.[15] The in-depth mechanics of the payment systems are outside the scope of this Section.[16] But, it is germane to mention that payment systems are broadly categorised into wholesale and retail payment systems, depending on the channel through which the funds are facilitated and the market participants.[17] Wholesale payments (and the linked securities settlement) involve the facilitation of high-value funds or assets transfer (using interbank arrangements). The Central bank acts as the coordinator of wholesale operations through its special status as the regulator of the banking system. Clearing and settlement of funds-involving the reconciliation of payments and disbursal of funds to payee respectively, is facilitated within the bank, through clearing houses or the central bank in the case of interbank payments within the domestic jurisdictions.[18]

On the other hand, retail payments involve the facilitation of low-value transfer of funds between individuals, households, businesses and government agencies. The core features of this category identified by Professor Hal Scott include: "universality (*i.e.* the ability to transfer funds at both point of sale and

[12] Ross Cranston, Emilios Avgouleas, Kristin van Zwieten, *et al*, *Principles of Banking Law* (3rd Ed, OUP 2017) 348.

[13] Financial Services (Banking Reform) Act 2013, section 41(2).

[14] Bruce Summers, *The Payment System: Design, Management, and Supervision* (International Monetary Fund 1994) ch 1 ("The Payment System in a Market Economy").

[15] Cranston, *et al*, *Principles of Banking Law* (n 12) ch 12; The International Bank for Reconstruction and Development / The World Bank, *Payment Systems Worldwide: A Snapshot. Outcomes of the Global Payment Systems Survey 2010* <http://siteresources.worldbank.org/FINANCIALSECTOR/Resources/282044-1323805522895/121534_text_corrections_5-7.pdf> (accessed 10 August 2018).

[16] For extensive reading of payment systems' nature and mechanics, see Cranston, *et al*, *Principles of Banking Law* (n 12) ch 12–13; Armour, *Principles of Financial Regulation* (n 11) ch 18; Awrey and van Zwieten (n 3).

[17] See Awrey and van Zwieten (n 3).

[18] Cranston, *et al*, *Principles of Banking Law* (n 12) ch 8.

remotely); ease of use (including widespread acceptance by merchants); certainty of payment (subject to some degree of payment reversibility for mistaken payments); liquidity; recordkeeping; safety and security; and financial inclusion".[19]

However, a distinctive feature of the retail payment systems is that it includes a high volume of low-value transactions that are processed (predominantly) by technology firms which provide financial services outside regulated banking perimeter (termed "non-banks"). These non-banks facilitate payments through the provision of overseas remittance, foreign exchange and provisions of innovative payment solutions-payment cards, electronic transfers and mobile payments.[20]

3. Core Functions of Payment systems

The key elements of payment systems as argued by Sheppard include authorisation and initiation of payments, transmission and exchange of payment instruction and settlement between participant banks.[21] These elements are made possible through other core obligations including the clearing and settlement of payments,[22] storage of funds (custodial and transactional) and liquidity. Awrey and van Zwieten argued that the core functions of payment systems (through TFIs in banking systems) relates to the provision of storage facilities of funds and the promise of liquidity upon demand.[23] In this Section, I shall focus on the two core functions—storage and liquidity—identified by the writers.

(i) Storage

A cursory look at the evolution of money confirms that it has existed in many forms including: barter, commodity (gold, metal, copper, iron), coin, paper, and recently, e-money.[24] A driver of this evolution, apart from the double coincidence of want, revolves around ease of storage and transportation of money.[25] This is a

[19] See Awrey and van Zwieten (n 3).

[20] *ibid*; Bank for International Settlements (Committee of Payments and Market Infrastructures), *Non-banks in retail payments* (September 2014) <https://www.bis.org/cpmi/publ/d118.pdf> (accessed 10 August 2018).

[21] David Sheppard, *Handbooks in Central Banking No. 8: Payment Systems* (Bank of England Centre for Banking Studies, London, 1996), cited in Michael C Blair, George Alexander Walker, Stuart Willey, *et al, Financial Markets and Exchanges Law* (2nd Ed, OUP 2012) 332.

[22] For clearing and settlement systems generally, see Cranston, *et al*, *Principles of Banking Law* (n 12) 349–52.

[23] Awrey and van Zwieten (n 3) 781–4.

[24] William Warrand Carlile, *The Evolution of Modern Money* (Macmillan and Co 1901).

[25] George Alexander Walker, Robert L Purves, and Michael C Blair, *Financial Services Law* (4th Ed, OUP 2018).

crucial function performed by TFIs—particularly banks—which collects deposits, store and provide funds upon demand for satisfaction of financial obligations.[26] Specifically to bank-based payment systems, the storage functions may be divided into custodial and transactional storage. Custodial storage concerns the protection of client funds (or payment instruments) from theft, fraud and destruction prior to its use for payment transactions. This was traditionally performed through storage in giant vaults and later, with the advent of technology, electronically in bank accounts and ledger balances to which clients have proprietary rights as secured depositors.[27]

On the other hand, transactional storage is performed by TFIs (banks and clearing houses) by facilitating the secure and efficient transfer of stored funds to third parties upon demand by owners for satisfying financial obligations.[28] With the advent of technology in the modern payment systems, these functions are performed using innovative institutional clearing and settlements systems—real time gross settlement (RGTS), deferred net settlements (DNS) and other jurisdiction specific automated systems.[29] Non-banks also provide similar services in collaboration with TFIs in offering "account-based" or "web-based" remittance and payment services to clients (particularly payment cards and mobile money).[30] The funds are stored in payment cards secured by Chip and Pin issued by the non-banks or mobile money wallets providers.[31]

(ii) Liquidity

In payment systems, liquidity concerns the availability of an asset in the required form for use in the purchase of goods and services without delay as to transferability or access.[32] It could also broadly be expressed to mean the timely redelivery of funds stored by payment service providers back to their owners upon demand for the purchase of goods or repayment of debt.[33] Banks are able to provide this function through the maintenance of a portion of their capital in

[26] Awrey and van Zwieten (n 3) 781–91.
[27] *ibid*.
[28] *ibid*.
[29] Armour, *Principles of Financial Regulation* (n 11) ch 18.2; Cranston, *et al*, *Principles of Banking Law* (n 12) 349–51.
[30] *Non-banks in retail payments* (September 2014) (n 20) 1.
[31] *ibid*.
[32] See Armour, *Principles of Financial Regulation* (n 11) footnote 17.
[33] Summers, *The Payment System* (n 14) 2.

liquid assets (Treasury bills) which may be easily disposed in the interbank market to provide funds in satisfaction of customers deposit withdrawals.[34]

Although the challenges to the bank's liquidity is outside the scope of this Section, it's worth mentioning that banking system is inherently unstable. The instability results from the banks use of short term deposits in financing long term loans resulting in a maturity mismatch.[35] To ensure the that this instability does not affect overall payment systems, central banks as the principal regulator of the banking and payment system, performs the role of lender of last resort and emergency liquidity provider to the banks where it suffers liquidity crisis.[36] This fund is made available to banks subject to strict preconditions of solvency, attendant systemic effects and at punitive interests to avoid the moral hazard of intentional exposures.[37] In the case of an institutional failure, depositors in TFIs are protected by deposit insurance and guarantee schemes which compensates depositors for any loss to savings.[38]

B. Virtual currencies

1. What are Virtual Currencies?

Defining virtual currencies (VCs) depends on whether the attempt is to consider it as a currency, an investment, or a payment network.[39] The paper's focus is on payment systems and I shall therefore limit myself to the consideration of VCs as a payment system (and as a payment instrument only where necessary). As a payment system, VCs can be defined simply as a peer-to-peer (P2P) operational network governed by rules and standards for transfer of electronic cash among members of a virtual community without need for financial intermediaries.[40] Not all VCs are payment systems, but the openly convertible VCs (Bitcoin) function

34 For banks liquidity requirements, see Cranston, *et al*, *Principles of Banking Law* (n 12) ch 2–3.

35 Douglas W Diamond and Philip Dybvig, "Bank Runs, Deposit Insurance, and Liquidity" (1983) The Journal of Political Economy 91(3) 401, 401.

36 Edward J Green, "The Role of the Central Bank in Payment Systems" in Andrew Haldane, Stephen Millard, and Victoria Saporta (eds), *The Future of Payment Systems* (Routledge: Abingdon 2007); Kern Alexander, Rahul Dhumale, and John Eatwell, *Global Governance of Financial Systems: The International Regulation of Systemic Risk* (OUP 2006) 184.

37 Thomas L Hogan, Linh Le, and Alexander William Salter, "Ben Bernanke and Bagehot's Rules" (2015) 47(2) Journal of Money, Credit and Banking 333–48.

38 Armour, *Principles of Financial Regulation* (n 11) ch 18.2.4.

39 For further readings on virtual currencies generally, see David Lee Kuo Chuen, *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data* (Academic Print, Elsevier 2015).

40 Satoshi Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System* (2008) <https://bitcoin.org/bitcoin.pdf> (accessed 9 August 2018); Hughes and Middlebrook (n 6) 517–8.

as payment systems through VC-intermediaries. Defining VCs as electronic cash is perhaps misleading, the primary reason is because, unlike electronic money, VCs are not expressed in the traditional fiat currencies of particular sovereign jurisdiction[41] or 'expressed in traditional accounting units, such as in Euro, but in virtual accounting units, such as the "bitcoin".[42] The system operates through an open source software (called "blockchain") accessible over the internet by all members of the virtual community and is used in the facilitation of VCs between members without the necessity of a trusted third-party intermediary.[43]

Since the creation of the first and most popular VC, Bitcoin in 2009 by Satoshi Nakamoto—an unknown software developer—several other VCs (termed 'altcoins') have been created using similar software protocol as bitcoin for their payment systems.[44] While some of these VCs operate in closed systems-with no convertibility to conventional fiat currencies, the focus of this Section is limited to freely convertible VCs (that is, VCs with values substitutable for fiat currencies)—particularly Bitcoin (termed broadly as "cryptocurrencies").[45] Bitcoin has the highest market capitalisation among the over 1700 VCs presently in existence and its freely convertible (through exchanges) to fiat currencies, thereby making it significant to traditional payment systems.[46] I will therefore briefly consider how it works and subsequently comment on it functioning as a payment system through Crypto-currency exchanges.

2. How Does Bitcoin Blockchain Work?

To understand how Bitcoin works as a payment method, it's prudent to briefly comment on its nature as a payment instrument, though outside the scope of this paper. Bitcoin may be defined as a private unregulated digital cash which is neither issued nor controlled by a sovereign institution (central bank) but created through special algorithms (cryptography) in a decentralised open distributed

[41] Armour, *Principles of Financial Regulation* (n 11) 370.

[42] *ibid* citing European Court of Justice, C-264/14 *Skatteverket v Hedqvist* [2016] STC 372 at [11].

[43] Brühl (n 2); Tracey Anderson, "Bitcoin-Is it a Fad? History, Current Status and Future of the Cyber-currency Revolution" JIBLR 428, 429; Nicholas A Plassaras, "Regulating Digital Currencies: Bringing Bitcoin within the Reach of the IMF" (2013) 14(1) Chicago Journal of International Law 377.

[44] For a comprehensive list of existing virtual currencies. see www.marketcap.com (accessed 10 August 2018).

[45] Financial Action Task Force, *Virtual Currencies: Key Definitions and Potential AML/CFT Risks* (June 2014) <https://www.fatf-gafi.org/media/fatf/documents/reports/Virtual-currency-key-definitions-and-potential-aml-cft-risks.pdf> (accessed 10 August 2018) 4–8.

[46] *ibid*.

network node (called 'mining') by its developers or special users (called miners).[47] Its distinguishing feature is pseudonymity of ownership because it exists essentially in virtual form capable of transfer through the internet using cryptography over a P2P network without regulation from any government authority or facilitation through a trusted TFI.[48]

For facilitating the transfer of value in bitcoin between the members of the virtual community (apart from crypto-currency exchanges platform), it uses an innovative payment protocol-Blockchain (a type of distributed ledger technology "DLT").[49] To initiate a transfer, the anonymous holder of the unique identifier number (called 'private key') effects a transfer of the agreed unit of the VC to a transferee's public key from the encrypted digital wallet ('bitcoin addresses') using an electronic signature. This generates a complex algorithmic problem with a timestamp on the transaction initiated—making it unalterable and irreversible.[50] While the timestamp is to prevent the likelihood of double spending or counterfeiting by the payor, the electronic signature and algorithmic puzzle is used to verify the ownership of the transferor and the validity of the initiated transactions.[51] The verification of the transaction is done by special users (called 'miners') using heavy computational protocols and cryptography in solving the complex mathematical problems generated by the transaction within 10–20 minutes through a 'proof of work'.[52] This verification ensures the integrity of the network by preventing internal/external fraud (by hackers) and encourage continuity by rewarding successful miners with newly created bitcoins. The verified transaction is thereafter entered as a block into the public network blockchain that is readily accessible by all members of the virtual community.[53]

Blockchain functions outside the traditional banking system using instead a P2P open network that can be accessed by the members of the virtual community at any time globally via the internet Web. Further, VCs operate within a self-regulatory framework-using cryptography and blockchain to prevent double-spending, fraud, and cyber-attacks (the adequacy of this measure is considered elsewhere in this paper). Although still significantly lower than other payment

[47] Cranston, *et al*, *Principles of Banking Law* (n 12) 369.

[48] See Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System* (n 40).

[49] Bitcoin, "How does Bitcoin work?" <https://bitcoin.org/en/how-it-works> (accessed 20 July 2018); Hughes and Middlebrook (n 6) 505.

[50] Cranston, *et al*, *Principles of Banking Law* (n 12) 370.

[51] *ibid*.

[52] Brühl (n 2) 371–3.

[53] *ibid*.

systems, the acceptability of VCs as payment system is growing progressively. This is despite the shunning by most regulators because it is unregulated and the bad press since its inception.[54] In the next Part, I shall examine whether VCs-through crypto-currency exchanges-perform core functions like TFIs in conventional payment systems.

C. Do virtual currencies perform core payment functions?

Earlier in this Section, I examined payment systems and the core functions performed by TFIs as PSPs. Two core functions—storage and liquidity—were considered using the banking system as focus. VCs, as evident from the facts in Part B, operates within a self-regulated open network outside the perimeter of regulated banking system without need for financial intermediaries in facilitating transfer of value. This financial independence was primarily born out of a distrust for TFIs which grew after 2008 Global Financial Crisis (GFC) that occasioned substantial losses to consumers.[55] This point, in addition to a craving for financial inclusion, was canvassed as the foundation for VC initiative in the white paper by Satoshi Nakamoto.[56] Attempts to avoid limitations and costs resulting from dependence on TFIs in payment processing heralded the era of 'shadow payment systems'—which are payment service providers who perform the core functions of traditional banks (deposit taking, storage and liquidity) outside the perimeter of the regulated banking system.[57] They are broadly classified into P2P payment systems (Paypal), mobile money platforms (M-Pesa) and Crypto-currency exchanges (Mt.Gox, Coinbase, CoinCheck).[58] This Section will focus on how Crypto-currency exchanges perform core payment functions like TFI.[59] I have chosen crypto-currency exchanges instead for individual VCs for two reasons: (a) there are presently over 1700 VCs; and (b) not all VCs are payment systems (including

[54] Cranston, *et al*, *Principles of Banking Law* (n 12) 370.

[55] For further readings on the global financial crisis, see Iain MacNeil and Justin O'Brien, *The Future of Financial Regulation* (Hart 2010); George A Walker, "Financial Crisis and Financial Resolution" (2013) 29(1) *Banking and Finance Law Review* 55.

[56] See Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System* (n 40).

[57] Awrey and van Zwieten (n 3) 796.

[58] *ibid* 777.

[59] For furthering readings on shadow payment systems generally, see Awrey and van Zwieten (n 3) Part V.

Bitcoin) but a significant number of VCs function as payment systems through Crypto-currency exchanges.[60]

Crypto-Currency Exchanges

Although Nakamoto argued against the use of financial intermediaries, most VCs transactions are facilitated through crypto-currency exchanges (hereinafter called 'the exchange') who perform similar functions as TFIs within the VC system.[61] Before its bankruptcy in 2014, Mt Gox.—founded in Tokyo in 2009—was reputed as the cornerstone of the Bitcoin system, facilitating more than 70% of the total bitcoin transactions globally.[62] In recent times, Coinbase (with other exchanges) collectively facilitate VC-transactions for more 30 million clients globally (including top retail merchants like PayPal, Microsoft and Amazon).[63] These exchanges operate outside the perimeter of the regulated banking system though they provide payment services like TFIs.[64] Their core services include: facilitating the interconvertibility between VCs and conventional fiat currencies (US Dollars, Euros, Pounds Sterling); providing online wallets on their servers for the storage of VCs to prevent theft and fraud and; providing a platform gateway for matching users in settling financial obligations.[65]

While exchanges facilitate the speedy transfer of value in VCs between users, their existence do not prejudice the ability of individuals to use the web-based system independently. It is the global network of the exchange's platform through which faster pairing of users is made possible that make their use attractive.[66] They provide storage facilities to users by issuing high security passwords to their clients for access to VCs deposited on the exchange's server (either stored online or offline).[67] To initiate a transfer of VC, the user need only request that a payment be effected in favour of another member of the virtual community using the specialised password provided upon registration.[68] Once matched, the exchange ensures the completion and settlement of the transaction. The record

[60] Hughes and Middlebrook (n 6) 517–8.
[61] *ibid*.
[62] Robert McMillan and Cade Metz, "The Rise and Fall of the World's Largest Bitcoin Exchange" WIRED (November 2013) <https://www.wired.com/2013/11/mtgox/> (accessed 10 August 2018).
[63] See Coinbase <https://www.coinbase.com/> (accessed 10 August 2018).
[64] Awrey and van Zwieten (n 3) 797; Hughes and Middlebrook (n 6) 495.
[65] Awrey and van Zwieten (n 3) 797.
[66] *ibid*.
[67] See Coinbase (n 63).
[68] Awrey and van Zwieten (n 3) 797–800.

may be added to the blockchain or off the block depending on whether both users are members of the same exchange.[69]

The second core function performed by exchanges like TFIs is the provision of liquidity upon demand. This function is performed by allowing the prompt conversion of VCs to fiat currencies through the server at prices determined by the forces of supply and demand between the VC and specific fiat currencies.[70] Where payment is made in VC for goods purchased from merchants, the exchange absolves the liquidity and currency exchange risks by its immediate obligation to pay the going exchange rate of the VC at the time of the transaction. This implies that if the value of the VC drops after the transaction, the merchant will still be entitled to the transactional value. Although the challenges to this function is not strictly the focus of this paper, it suffices to mention that exchange users are regarded as unsecured creditors unlike TFIs and have no first claim during bankruptcy.[71] This was what happened with Mt. Gox in 2014 after it suffered losses to assets (worth $470m) through cyber-attacks.[72] The exchange's operation outside regulated banking system implies that it does not receive support of emergency liquidity or lender of last resort from central banks like TFIs. It also does not benefit from deposit insurance schemes or guarantees.[73] While these concerns may cripple its functioning, it is without doubt that, except where an institutional failure occurs, VCs function as payment systems through crypto-currency exchanges. The operational network and standardised transaction rules are enforced by the exchanges and like TFIs, they perform core functions of payment systems within the financial systems.

I have established in this Section that VCs—through crypto-currency exchanges—perform similar functions as the TFIs in conventional payment systems. In the next Section, I will focus more on the indirect interdependencies of global payment systems through the activities of multinational TFIs. I will also

[69] *ibid*.
[70] *ibid*.
[71] *ibid* 799.
[72] *ibid* (footnote 132).
[73] *ibid*.

examine the impact a collaboration between a crypto-currency exchange and a TFI might have on these interdependencies.

III. Impact of VC exchanges on global payment systems

The aim of the Section is to critically analyse how VC intermediaries may pose potential disruptions to the global payment and settlement systems. The Section is divided in two parts. Part A will examine how the global payment system works with a focus on financial consolidation and technological innovation as major drivers of the interdependencies in transnational payment systems.

In Part B, I will attempt the question of whether exchanges may pose potential disruptions to the global payment systems through a possible financial integration with key stakeholders (TFIs) in the global payment network. In this part, I will employ a hypothetical situation in which VC exchanges collaborate with multinational TFIs to consider the effects such integration may have on the global payment system. My reasons for choosing financial consolidation over other factors are twofold: (a) Recent trend in global systems confirms large TFIs are considering the adoption of blockchain (though a modified permissioned closed systems format) as an alternative to the costly and time consuming processing of cross-border payments using traditional systems; and (b) The initiative for the creation of Central Bank Digital Currencies (CBDC)as an alternative payment instrument is gaining momentum among major central banks globally.[74] These two initiatives suggest the likelihood of a collaboration between VC exchanges and TFIs soon if adopted.

A. How global payment systems work

Due to the transnational scope of VCs and exchanges, I have opted to focus on the role of TFIs in the taxonomy of cross-border payment systems. However, the global payment system comprises a network of both domestic and cross-border systems which are interdependent in achieving the efficient flow of funds among global financial systems.[75] The network comprises multinational TFIs (operating

[74] Bank for International Settlements (Committee on Payments and Market Infrastructures and Markets Committee), *Central bank digital currencies* (March 2018) <https://www.bis.org/cpmi/publ/d174.pdf> (accessed 6 July 2018).

[75] See Bank for International Settlements (Committee on Payment and Settlement Systems), *The Interdependencies of Payment and Settlement Systems* (June 2008) <https://www.bis.org/cpmi/publ/d84.pdf> (accessed 10 August 2018); *Payment Systems Worldwide* (n 15). For further readings on domestic payment systems, see Cranston *et al*, *Principles of Banking Law* (n 12) ch 13.

directly or as 'correspondents or custodians'), central banks, other PSPs and central clearing depositories as its key participants.[76] Interdependence of financial systems may be system-based or institution-based depending on whether the relationship is direct or indirect. Direct relationships arise from the use of the same payment messaging service providers, central clearing depositories, payment processors and risk management.[77] On the other hand, indirect interdependencies emanates predominantly from the activities of Large TFIs (termed globally systemically important financial institutions, "GSIFIs")[78] or PSPs operating within multiple jurisdictions and indirectly linking the stability of each financial system to the smooth running of the others in which they operate.[79]

The crucial drivers of the evolution of interdependencies among financial systems identified by the CPSS (now CPMI) include: globalisation, trade liberalisation; financial consolidation; regional integration; technological innovation, public policies.[80] This list has expanded to include E-commerce and mobile telecommunications in recent times.[81] Cross-border wholesale payments (termed 'Systemically important payment systems')[82] which initially involved the use of traditional legacy-based payment methods-documentary credits,[83] are now predominantly facilitated through modern 'interbank' payment methods.[84] These drivers also influenced the dynamism of retail payment systems, making them significant processors for non-cash transnational payments processing.[85] Modern payment methods provided by TFIs and non-banks-payment cards, electronic

76 *ibid*.

77 See *Interdependencies of Payment and Settlement Systems* (*ibid*) 1–5; *Payment Systems Worldwide* (n 15) 81.

78 Financial Stability Board, *Policy Measures to Address Systemically Important Financial Institutions* (November 2011) <https://www.fsb.org/wp-content/uploads/Policy-Measures-to-Address-Systemically-Important-Financial-Institutions.pdf> (accessed 10 August 2018).

79 *ibid* 2.

80 Chris Brummer, *Soft Law and the Global Financial System: Rule Making in the 21st Century* (CUP 2012) 10.

81 Bank for International Settlements (Committee on Payment and Settlement Systems), *Innovations in Retail Payments* (May 2012) <https://www.bis.org/cpmi/publ/d102.pdf> (accessed 10 August 2018); for further readings on E-commerce, see Paul Todd, *E-Commerce Law* (Cavendish 2005).

82 Bank for International Settlements (Committee on Payment and Settlement Systems), *Core Principles for Systemically Important Payment Systems* (January, 2001) <https://www.bis.org/cpmi/publ/d43.pdf> (accessed 10 August 2018).

83 Alastair Hudson, *The Law of Finance* (Sweet & Maxwell 2013) ch 30.

84 Awrey and van Zwieten (n 3) 791; Bank for International Settlements (Committee on Payment and Settlement Systems), *New Developments in Large-Value Payments* (May 2005) <https://www.bis.org/cpmi/publ/d67.pdf> (accessed 10 August 2018).

85 European Central Bank, "Retail Payments and the Real Economy" Working Paper No. 1572 (August 2013) <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1572.pdf?0568b27871896eb01f54b0c4c40a8f63> (accessed 10 August 2018).

fund transfers, mobile money and web-based payments have since become the preferred alternatives for domestic and cross-border payment transactions.[86]

Apart from technological innovation, financial consolidation (particularly mergers and joint ownership) has been the major driver of indirect interdependence of global payment systems.[87] TFIs are increasingly collaborating among themselves and with other entities: financial technology firms (Fintech), Central depositories and risk management providers in becoming an all-round financial services provider in the global financial system.[88] This has influenced the volume and value of non-cash payment transactions (over $400 trillion in 2014) facilitated within the global payment system and the increase in global gross domestic product (GDP).[89]

While wholesale large-value systems (facilitated by TFIs) are considered systemically important payment systems because of the value of transactions processed, the retail payment systems have also witnessed an upward review in recent times.[90] This shift predominantly resulted from financial consolidations between TFIs and new market entrants into the payment industry—Non-banks (or Fast payment schemes).[91] Non-banks (and Fast payments) are institutions which provide payment services (and other financial functions) outside the perimeter of the regulated banking system with the capacity to process payments and settlement at any time of the day in (near) real-time.[92] The major drivers of the integration are profit maximisation and regulatory arbitrage.[93] Profit maximisation results from the economies of scale and scope that comes with the expansion of business and efficiency associated with the use of technology in facilitating payment processing and settlement. Arbitrage—the legal avoidance of strict compliance with regulations— is an incentive to the TFIs because technology firms (non-banks) provide payment services outside the complex regulatory perimeter of TFIs.[94] By

[86] Awrey and van Zwieten (n 3) 800–2.

[87] For further research on Financial Consolidation and Corporate Restructuring see Patrick A Gaughan, *Mergers, Acquisitions, and Corporate Restructurings* (6th Ed, Wiley 2015).

[88] *Interdependencies of Payment and Settlement Systems* (n 75) section 2.

[89] Awrey and van Zwieten (n 3) footnote 1.

[90] *Innovations in Retail Payments* (n 81) Introduction.

[91] "fast payment" is defined as a payment in which the transmission of the payment message and the availability of "final" funds to the payee occur in real time or near-real time on as near to a 24-hour and seven-day (24/7) basis as possible: see *Fast Payments – Enhancing the Speed and Availability of Retail Payments* (n 5) section 2.1.

[92] Bank for International Settlements (Committee on Payments and Market Infrastructures), *Non-Banks in Retail Payments* (September 2014) 1; *Fast Payments – Enhancing the Speed and Availability of Retail Payments* (n 5) Executive Summary.

[93] For further readings on Arbitrage, see Joanna Benjamin, *Financial Law* (OUP 2007) ch 23.

[94] *Non-Banks in Retail Payments* (n 92) section 3.

extension, services provided by TFIs through the non-banks are not subjected to the strict regulations and attendant operational costs.[95]

Perhaps a more important question is how financial consolidation drives the interdependence of global payment (and financial) systems generally? A CPSS survey suggests that new payment systems are based majorly within domestic jurisdictions with several existing globally.[96] The report further suggests that these new payment systems (non-banks, P2P systems and mobile money) benefit from cross-border transactions through cooperation and standardisation among themselves and through collaborations with TFIs.[97]

These revelations have two crucial impacts on the global payment systems. The first and most evident is that it creates an efficient system of interlinked financial systems and institutions within the global payment system.[98] The collaborations between TFIs and non-banks (through joint ownership or outsourced payment and technology related services) ultimately creates an indirect interdependence between several financial systems in which the institutions operate thereby facilitating the speedy processing of transactions within the global network.[99] The second and arguably negative consequence of these integration is that it creates a form of 'systemic risk' between the interdependent financial systems.[100] Major risks of payment systems include: internal/external fraud, employee misconduct and model design collapse (termed 'operational risk'); inability to complete transaction due to lack of immediate funds (called 'liquidity risk' or 'payment and settlement risk'); institutional failure (broadly termed 'market risk')and ;credit risks from time delays.[101] Consolidation among TFIs (and new market entrants) results in exposure

95 *ibid*.

96 *Innovations in Retail Payments* (n 81) Executive Summary.

97 *ibid* section 4.2.

98 *Interdependencies of Payment and Settlement Systems* (n 75) Part II.

99 *ibid*.

100 Systemic risk is defined as "the probability that cumulative losses will occur from an event that ignites a series of successive losses along a chain of [financial] institutions or markets comprising... a system": see Steven L Schwarcz, "Systemic Risk" (2008) 97(1) Georgetown Law Journal 193, footnote 11; G Afonso and HS Shin, "Systemic Risk and Liquidity in Payment Systems" (2009) Federal Reserve Bank of New York Staff Report No 352. For further definitions and readings on systemic risks, see: Alexander *et al*, *Global Governance of Financial Systems* (n 36) 23–4 and ch 7.

101 *Interdependencies of Payment and Settlement Systems* (n 75) Part II; Alexander *et al*, *Global Governance of Financial Systems* (n 36) 188–90.

by one financial institution to the risks particular to other participants due to their interconnectedness.[102]

Thus, failure of either of the interconnected institutions could pose systemic risks to the consumers, market participants. These failures may also extend to the several national economies in which the affected institutions operate because of the crucial role payment systems play.[103] For example, the institutional failure (through license withdrawal) of the Bankhaus Herstatt—an international financial institution— in 1974 resulted in settlement risk to the foreign exchange market in which it was an active participant.[104] The failure of this GSIFI affected finality of transactions facilitated through it. Thus, triggering systemic failures not only in the financial institutions connected to the German bank but the financial systems in which these affected institutions were key stakeholders of financial stability.[105]

From facts and examples briefly considered above, its discernible that the global payment systems function through the interdependence among the financial institutions and systems which are key stakeholder in the payment systems network. These interdependencies have the advantages of increased efficiency, economies of scale and scope and better risk management among participating systems and institutions due to broader risk sharing and risk management.[106] However, these advantages may be whittled down considerably if one considers the impact the failure of any of the interdependent institutions or systems may have on other participants or even the global financial system generally. The role played by TFIs is also evidently crucial to the network. In the next Part, I will examine how VC exchanges may impact the global payment systems (positively or otherwise) if (and when) it becomes a key stakeholder (by consolidation) in the global payment and settlement systems.

B. How VC exchanges may impact global payment systems

Despite VC-blockchain's lack of upper limit to the value of funds that may be processed through the network, the system is used for low-value transactions

[102] *Interdependencies of Payment and Settlement Systems* (n 75) section 5; Bank for International Settlements (Committee on Payment and Settlement Systems), *Cross-Border Securities Settlements* (March 1995) <https://www.bis.org/cpmi/publ/d12.pdf> (accessed 10 August 2018); Bank for International Settlements (Committee on Payment and Settlement Systems), *The Role of Central Bank Money in Payment Systems* (August 2003) <https://www.bis.org/cpmi/publ/d55.pdf> (accessed 10 August 2018).

[103] *Interdependencies of Payment and Settlement Systems* (n 75) 26.

[104] Armour, *Principles of Financial Regulation* (n 11) 399. For research on similar failures, see also: Dominique Rambure and Alec Nacamuli, *Payment Systems: From the Salt Mines to the Board Room* (Basingstoke: Palgrave 2008).

[105] Armour, *Principles of Financial Regulation* (n 11) 399.

[106] *Interdependencies of Payment and Settlement Systems* (n 75) section 5.1.

presently.[107] This prevents it from being considered as systemically important to the global payment systems.[108] Perhaps, this is because VC exchanges are still relative new and do not benefit from the goodwill enjoyed by TFIs.[109] The European Banking Authority in a recent report concludes that VCs and its exchanges are predominantly used by private individuals with little institutional participation which drives the significance of TFIs as systemically important to wholesale payments systems.[110] This may not come as a surprise considering that VCs systems were designed to facilitate payments without need for trusted third party or financial intermediaries. But, considering the focus of these Section, the question could be asked that what impact can VC exchanges have on the interdependent global payment systems if they integrate with TFIs? To answer this question, I will assume that the VC exchanges become integrated with TFIs in the banking system—a key stakeholder in the global payment system. Financial consolidation occurs when two financial institutions performing similar or complementary functions come together under any form of corporate restructuring.[111] Put simply, what impact would a consolidation between a major bank (for example, JP Morgan Chase, Goldman Sachs) and a VC exchange have on the global payment system?

Like the drivers collaborations between TFIs and Non-banks identified by the Committee on Payment Markets Infrastructure(CPMI), the motivators of a consolidation between TFIs (which indirectly connect the global payment system) revolves around the benefit of technological innovation.[112] Through consolidation, ancillary benefits such as profit maximisation from economies of scale, lower transactional and regulatory costs and increased efficiency become available to TFIs.[113] For VC exchanges, the most important driver will likely be the increased trust and integrity the consolidation will bring to its operations.[114] However, this consolidation may also create a pathway for VC exchanges direct link to the global payment system. This is because TFIs in banking systems indirectly connect

107 See Blockchain, "Blockchain Size" <https://www.blockchain.com/en/charts/blocks-size> (accessed 7 August 2018).

108 European Central Bank, *Virtual Currency Schemes* (October 2012) <https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf> (accessed 7 August 2018).

109 Financial Stability Board, *Letter to G20 Ministers and Central Bank Governors* (18 March 2018) <https://www.fsb.org/wp-content/uploads/P180318.pdf> (accessed 7 August 2018).

110 European Banking Authority, *Opinion on Virtual Currencies* (July 2014) <https://eba.europa.eu/sites/default/documents/files/documents/10180/657547/81409b94-4222-45d7-ba3b-7deb5863ab57/EBA-Op-2014-08%20Opinion%20on%20Virtual%20Currencies.pdf?retry=1> (accessed 7 August 2018) Introduction.

111 See Gaughan, *Mergers, Acquisitions, and Corporate Restructurings* (n 87).

112 *Non-Banks in Retail Payments* (n 92) section 3.

113 See *Interdependencies of Payment and Settlement Systems* (n 75); *Virtual Currencies and Beyond* (n 4) 18–20.

114 Hughes and Middlebrook (n 6) 499, footnote 9.
Awrey and van Zwieten (n 3) 797.

several national payment systems through their activities globally. Emerging issues of TFIs modern consolidations with new entrants (particularly non-banks) revolve around exposures to operational risks (cyber security, identity theft and fraud), liquidity risks and prudential regulation of new market entrants.[115] To combat these issues, several measures including: Chip and Pin security and Know-Your-Customer (KYC) requirements have been created by regulators to ensure safety.[116] The efficiency of these measures are beyond the scope of this paper.

Having mentioned the drivers of a possible consolidation between TFIs and VC exchanges, the likely positive impact the integration will have on global payment systems may include: increased efficiency, financial inclusion-to the underbanked national economies with limited access to TFIs, and economies of scale to market participants from the speed and relatively reduced cost associated with use of blockchain as an innovate payment solution.[117]

In the next Section, the more important question of what impact an interdependence between TFIs and VC-exchanges may have on the global payment system stability is examined. The immediate and extended effects of an exposure by the TFI to VC exchange's operational risk-cyber security is analysed critically.

Implications of Cross-Pollinating VC Exchanges' Risks to Global Payment System

Despite the several risks of VCs generally, the most systemic risks of VCs and their exchanges revolves around cyber-security.[118] Although not peculiar to the VC exchanges, the frequency with which cyber-attacks have ravaged the exchanges since their inception has generated concerns among domestic and international financial systems' regulators.[119] These attacks are from cyber-criminals who hack into the exchange's platform and cause theft of VCs through surreptitious mining activities. The vulnerability of the exchanges to attacks are primarily from the features of VCs including: anonymity, lack of centralised authority (regulator) and

[115] Armour, *Principles of Financial Regulation* (n 11) 407–8; Bank for International Settlements (Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions), *Principles for Financial Market Infrastructures* (April 2012) <https://www.bis.org/cpmi/publ/d101a.pdf> (accessed 9 August 2018).

[116] Awrey and van Zwieten (n 3) 796; Armour, *Principles of Financial Regulation* (n 11) 407.

[117] *Interdependencies of Payment and Settlement Systems* (n 75) 14.

[118] Bank for International Settlements (Committee on Payments and Market Infrastructures), *Digital Currencies* (November 2015) <https://www.bis.org/cpmi/publ/d137.pdf> (accessed 9 August 2018); *Opinion on Virtual Currencies* (n 110) 21–36; Lawrence J Trautman and Alvin C Harrell, "Bitcoin versus Regulated Payment Systems: What Gives" (2017) 38(3) Cardozo L Rev 1041.

[119] *Opinion on Virtual Currencies* (n 110) 44; and generally, *Virtual Currencies and Beyond* (n 4).

irreversibility of transactions.[120] Since the Mt. Gox bankruptcy in 2014, several cyber-attacks involving millions in stolen VCs have ravaged subsequent exchanges with the recent in CoinCheck where $500m worth of Bitcoins were stolen from the exchange's online wallet.[121] These attacks pose systemic risks to the entire virtual community presently but not the financial systems because they are not interlinked with the banking system TFIs.[122]

Thus, a consolidation between TFIs and VC exchanges may result in the cross-pollination of risks of cyber-attacks and liquidity crisis through their interdependence. This will likely emanate from TFI's use of VC exchanges for the processing of high-value cross-border payment instructions due to the blockchain's speed, reduced costs and efficiency in clearing and settlement of payment transactions. An example could be made of the likely effect a loss (through cyber-attacks) of a high-value cross-border payment transaction jointly processed by a TFI and VC exchange will have principally on the banking system as a key stakeholder in the global payment network and the 'knock-on effect' the banking systems failure will have on financial systems and institutions interdependent on it. This indirect interdependence through the activities of GSIFIs results in a new breed of "cross-system" risks— in which the failure of a financial system or institution could have a domino effect on all other systems that are dependent upon its continuity for their stability.[123]

Apart from the Bankhaus Herstatt failure earlier used as an example in this Section, a closely related example could be made of the effect the failure of Lehman Brothers—another multinational TFI— in the United States had on all other financial institutions and national economies linked to the US Subprime Mortgage market.[124] Admittedly, this example does not concern payment systems, but it serves as a reference point for the increased interdependence of the global financial systems. It also confirms the domino effect the failure of a key stakeholder can have on other market participants, consumers and national economies interdependent on it.[125]

The examples considered above suggests that VC exchanges effect on global payment system could result in increased efficiency and a possible paradigm shift in

[120] *Virtual Currencies and Beyond* (n 4) 24–33

[121] See Cointelegraph, "How to deal with Cyber-Attacks" (March 2019) <https://cointelegraph.com/news/mt-gox-coincheck-binance-and-more-how-exchanges-are-learning-to-deal-with-cyber-attacks> (accessed 9 August 2018).

[122] *Virtual Currency Schemes* (n 108) 39–40.

[123] *Interdependencies of Payment and Settlement Systems* (n 75) 28–30.

[124] Mark T Williams, *Uncontrolled Risk: The Lessons of Lehman Brothers and how Systemic Risk can Still Bring Down the World Financial System* (McGraw-Hill 2010).

[125] *ibid*.

the global payment infrastructure. These effects will stem from an increased volume in the number large-value payments across several financial systems through TFIs collaborating with VC exchanges. The blockchain open distributed network may also eliminate payment and settlement risks associated with cross-border payments while encouraging financial inclusion and transparency since transactions are concluded in (near)real-time and recorded on the block.[126] On the flip side, the consolidation could result in additional exposures by TFIs (and banking system) to the operational risks (particularly cyber security) of VC exchanges if not managed.

While this risk may appear minimal presently, it may quickly pose considerable risks if the desire for increased profitability by TFIs occasion an increased use of the VC exchanges for high value transactions. If this happens, the failure of any such transaction from exposures (to which VC exchanges are vulnerable) could cause potential disruptions to the reliant global payment systems. This could further result in systemic failures of interdependent financial systems (their consumers and market participants) who rely on each other for payment processing, risk management, liquidity and settlement in transnational flow of financial assets.[127] Disruptions from indirect relations are known to creep spread within interdependent networks through complex paths with uncertain levels of intensity.[128]

For regulators of payment systems, the emerging issues from this VC-TFI consolidation will primarily concern the role of Central banks as provider of liquidity and prudential oversight for the domestic payment systems.[129] The use unregulated VC exchanges pose two major challenges to regulators. First, stability of the banking system (and payment system) is maintained by central banks through the provisions of emergency liquidity and general monetary and economic policies.[130] A consolidation between TFIs and VC exchanges which operate globally may affect the functioning of Central banks. This challenge and possible solutions are dealt with later in this paper.[131] Lastly, VC exchanges operate a predominantly self-regulatory regime (using cryptography and blockchain) outside the perimeter of the regulated banking system despite its provision of core financial functions. The anonymity of its transactions implies that its popularity

[126] For payment and settlement risks, see *Interdependencies of Payment and Settlement Systems* (n 75) 27; Alexander *et al*, *Global Governance of Financial Systems* (n 36) ch 7.

[127] *Interdependencies of Payment and Settlement Systems* (n 75) 31.

[128] *ibid* 42.

[129] *Virtual Currency Schemes* (n 108) 33–42.

[130] *ibid*.

[131] *Non-Banks in Retail Payments* (n 92) section 6.

could have serious consumer protection issues that are beyond the scope of the users and (possible the central banks) as regulators.[132]

In concluding this Part, it is arguable from empirical evidence that TFIs adopt innovative payment solutions in the form of collaboration with modern technology providers for the provision of financial services with minimal operational costs.[133] Although the banking system is not directly linked to VC exchanges presently, the evidence of initiatives by TFIs on adopting blockchain suggests the likelihood of a future alliance.[134] If this trend gains popularity among financial systems and TFIs who are key stakeholders in the global payment systems, it could potentially set in motion a form of interdependence in which VC exchanges are stakeholders. This could ultimately result in potential disruptions and systemic risks to the entire global payment system if any institutional failure occurs to VC exchanges considering the interdependence it would have with TFIs. The primary barrier to the popularity of VC wide acceptance of VC exchanges emanates from its operational risks—particularly cyber security— in addition to its unregulated, anonymous and irreversible payment operations.[135] These challenges are not uncommon to technological innovations (as can be seen from the activities of other financial technology PSPs) and can be tackled using the technology and coordination.

From the above reasons, I argue that a financial consolidation between a VC exchange and a TFI could make the global payment system more efficient and cost friendly. However, the collaborations could also pose potential disruptions and systemic risks to the interdependent global payment systems if the exposures of VC exchanges are left unchecked. A possible cross-pollination of risks could result in the future if the exchanges become widely accepted -particularly in high value payments. This conclusion is admittedly speculative considering the likelihood of such consolidation is dependent on the actualisation and success of several proposed initiatives. Still, it provides an insight into the implications such integration could breed.

In the next Section, having established that VC exchanges could advance the functioning of global payment systems if the operational risks are better

132 See *Non-Banks in Retail Payments* (n 92): Drawing inference from the similarity in the mechanisms of Virtual currency systems and non-banks operating outside the regulated perimeter of banking system.

133 *ibid* Executive Summary.

134 See Blockchain, "Major Banks Have adopted Blockchain" (November 2017) <https://blockchain.works-hub.com/learn/Which-Major-Banks-Have-Adopted-or-Are-Adopting-the-Blockchain-> (accessed 9 August 2018.

135 *Virtual Currency Schemes* (n 108) 42.

managed, I would examine whether VCs and their exchanges require regulatory interventions like TFIs and more importantly, by who?

IV. Who should regulate VC payment systems?

In the preceding Section III, I examined the benefits and potential disruptions-systemic risks-that VCs exchanges may pose to the global payment system through a possible financial consolidation with TFIs. The major exposure of VC exchanges identified revolves around cyber security primarily resulting from anonymity and irreversibility in VC dynamics. In this Section, my aim is to examine whether VC exchanges require regulatory interventions from central authorities like TFIs in payment systems? If yes, By who? To achieve this objective, the Section is divided into three parts.

Part A will briefly consider the desirability of regulatory intervention on VCs exchanges. In Part B, focus will be on the regulators of TFIs in traditional payment systems. For this part, I shall limit my scope to central banks as the major regulator. Their limitations on extraterritoriality will also be analysed. Lastly, Part C will consider the desirability of regulating VC exchanges through international regulatory bodies in collaboration with domestic regulators. I should state presently that the aim of this Section is not to provide an exhaustive research on the subject but merely to explore some conceptual legal alternatives that may be referenced for further research.

A. Do VC exchanges require regulatory intervention?

As mentioned in Section II, VCs operate a self-regulated system in which all members of its virtual community (especially its developers) are responsible for maintaining the safety and integrity through their activities on the blockchain.[136] VC exchange system are especially designed this way to ensure financial inclusion and independence by avoiding the regulatory costs and limitations that plague TFIs.[137] Studies suggests that while VC exchange's system is essentially technology driven, its self-regulatory approach is not entirely new. Similar practice was recorded in the 19th century under the 'Suffolk banking system' where payment settlement and clearing is done 'in house'.[138] But, the regulation of any payment system must

[136] Awrey and van Zwieten (n 3) footnote 141.

[137] *ibid* footnote 140.

[138] *ibid* 790, citing Charles Calomiris and Charles Kahn, "The Efficiency of Self-Regulated Payment Systems: Learning from the Suffolk System" (1996) Journal of Money, Credit and Banking 766, 780; Gary Gorton, "Free Banking, Wildcat Banking and the Market for Bank Notes" (Wharton Financial Department, Working Paper, 1989).

critically focus on its efficiency, convenience, and safety.[139] While blockchain web-based system appears efficient and convenient for its continuous availability, the continuous cyber-attacks that have plagued VC exchanges suggests that the safety of the system leaves much to be desired presently.

Whether this vulnerability is because of a flaw in the blockchain system or internal negligence of the crypto-currency exchanges is beyond the scope of this paper. Suffice to reiterate that an emerging issue concerning VC exchanges (like other 'shadow payment systems') revolves around their capacity and incentive to take adequate measures to assess and manage the risks to which their systems (and their consumers) are exposed.[140] For example, during the notorious Mt.Gox bankruptcy, the users of its platforms were not protected by any form of insurance for their deposits and where therefore regarded as unsecured creditors under bankruptcy laws.[141] In this example, it appears that the losses suffered by the consumers using the Mt.Gox platforms was as a result of lack of prudential regulation guiding the internal structure and functioning of the exchange. To prevent the occurrence of this risk on its platform, Coinbase stores only 20% of its VCs holdings in the online wallets while the rest are stored offline.[142] While this measure might ensure better efficiency to custodial storage, it leaves open the window of possible attacks during the use of the platform as gateways for online virtual payments. To remedy the exposures in the VC exchanges, Awrey and van Zwieten have proposed some strategies including: storage of deposits with TFIs (Piggy-banking), deposit insurance with third-party schemes and the holding of deposits by VC exchanges as trusts.[143] While these strategies are plausible, their efficiency in the event of a financial consolidation between VC exchanges and TFIs cannot be determined conclusively. This is because the laws attempting to regulate VCs are divergent and uncertain presently.

The VC-TFI consolidation, as earlier mentioned, may pose significant systemic risks to global payment systems interdependent on it.[144] With the likelihood of VC exchanges facilitating transnational large-value payments through the conglomerate, special attention must be given to ensuring the safety of the system to avoid disruptions and externalities its failure might result on consumers, market participants and the global payment systems.[145] Reputational risks to the integrity

[139] See *Non-Banks in Retail Payments* (n 92).

[140] Armour, *Principles of Financial Regulation* (n 11) 407.

[141] Awrey and van Zwieten (n 3) 809.

[142] See Coinbase <https://www.coinbase.com/> (accessed 12 August 2018).

[143] Awrey and van Zwieten (n 3) 808–816. For measures to virtual currencies generally, see *Opinion on Virtual Currencies* (n 110) 38–45.

[144] *Opinion on Virtual Currencies* (n 110) 35.

[145] *Interdependencies of Payment and Settlement Systems* (n 75).

of the whole global payment systems from a failure of VC exchange also makes it imprudent to leave the regulation of such a potentially significant institution in the hands of "unknown regulators".[146] As a potential systemically important payment system (by association with the TFIs), VC exchanges require adequate rules on prudential regulation to ensure best practices within its internal and external functioning.[147] These regulations should naturally be from authorities with vast expertise and experience in financial dynamism and risk management which is hardly possessed by the software developers in which VC exchange's control are presently vested.[148]

Discernible from the facts considered above is that while VC exchange self-regulatory regime might appear efficient for miniature payment systems, it is grossly inadequate where it involves a broader interdependent network of financial systems.[149] The oversight by an external experienced regulatory body (for instance, a central bank) will not only give credence to the integrity of the payment system as a viable alternative, it would also opportunity to study the taxonomy and better manage possible risks from its operations.

In the next Part, I will examine the role of central banks as the major regulator of traditional payment systems. Its limitations as to scope of authority and efficiency in cross-border payments will also be analysed. The purpose of this examination is to tease out the challenges of the present divergent attempts by national regulators to regulate the global operations of VC exchanges.

B. Who regulates traditional payment systems?

The question of who regulates payment systems, TFIs and PSPs within national economies is not hard to answer. This is because, as mentioned in Section II, payments facilitation is among the core financial functions performed by traditional deposit-taking banks.[150] By implication, they are regulated alongside the traditional banking system. Other regulators of the securities system (including the Securities and Exchange Commission) also play significant role in the regulation of payment systems although through collaboration with the banking system.[151] The major regulators of the banking system (and by extension, payment systems) within domestic jurisdictions are the central banks.[152] For example, the Bank of

[146] Hughes and Middlebrook (n 6).
[147] *ibid*.
[148] *ibid*.
[149] For different virtual currency schemes, see *Virtual Currency Schemes* (n 108).
[150] Awrey and van Zwieten (n 3).
[151] Armour, *Principles of Financial Regulation* (n 11) section 18.3.3.
[152] Green, "The Role of the Central Bank in Payment Systems" (n 36); *Payment Systems Worldwide* (n 15) 1.

England in the United Kingdom, the Federal Reserve in the United States and regional European Central Bank in the Eurozone.[153] The central bank is specially crafted as a regulator of payment systems through its role as a catalyst and liquidity provider.[154]

As a catalyst, the central bank oversees the activities of TFIs and ensures stability of the payment systems through regulations, rules and guidelines reflecting its risk reduction policies.[155] This could be in the form of extended monetary policies on the flow of financial instruments (particularly fiat currencies) and broader prudential regulations on activities of key players within TFIs. This prevents operational risks that may result in institutional failures and negative effects on interdependent payment networks and markets.[156] As a liquidity provider, the central bank regulates the stability of payment systems by providing liquidity gap-filling through emergency intra-day credit facilities to TFIs(banks). In extreme cases that may threaten stability of financial systems, it acts as a lender of last resort in addition to its special resolution regimes on, market entry and exit of TFIs.[157]

This role played by central banks is crucial to the overall functioning and stability of the payment system within the national economies. In cross-border payments—involving several TFIs within different jurisdictions—the different central banks act as the central clearing institutions for the payments facilitated through their TFIs. For this role, they employ automated international payment messaging systems (for example, SWIFT) and foreign exchange clearing and settlement system-continuous linked settlements (CLS).[158] While regulatory policies and objectives are clearer in domestic payment systems, issues arise with their implementation on TFIs which operate within multiple jurisdictions.[159] Critical queries which challenge efficient implementation include: Which of the central banks in the different jurisdictions should regulate the TFI? Should it be the central bank in the home country where the TFI's headquarters is situated? Or better still should every central bank treat each subsidiary within its jurisdiction as a distinct entity and Brummer regulate it as such?

These questions are important for two major reasons: (a) central banks are established pursuant to the legislative instruments of sovereign jurisdictions with clearly defined powers and scope. They therefore (in most cases) lack the powers to regulate TFIs operations outside their territory because it poses threat to

[153] Hughes and Middlebrook (n 6) Part II.
[154] Alexander *et al*, *Global Governance of Financial Systems* (n 36) 184.
[155] *ibid*.
[156] Awrey and van Zwieten (n 3).
[157] *ibid*.
[158] Armour, *Principles of Financial Regulation* (n 11) section 18.2.3.
[159] Brummer, *Soft Law and the Global Financial System* (n 80) ch 1.22.

sovereignty of other jurisdictions; and (b) since cross-border large-value payments are majorly facilitated through multinational TFIs, the risk of systemic risks crossing territorial borders from through a TFI subsidiary to other branches in the financial conglomerate (and the financial systems in which they operate) increases exponentially.[160]

Consequently, although a central bank may efficiently manage risks within its jurisdiction, it remains permanently exposed to systemic risks from other jurisdictions due to its indirect interdependence through reliance on activities of TFIs operating in multiple jurisdictions.[161] The concerns are particularly exacerbated by the fact that central banks—as a limb of the sovereign authority—have diverse approaches to how TFIs and financial systems should be regulated.[162] While some financial systems, for example the United States, are heavily regulated, others—like Switzerland—operate a loosely regulated financial system. This disparity and competition among central banks results in regulatory arbitrage—through which financial institutions fashion payments within regulatory gaps for profit maximisation.[163]

Drawing inference from the above, are central banks capable of regulating VC exchanges alone? The answer is arguably, 'no'. This is because, like Mt.Gox, most exchanges, although located within national jurisdictions, operate significantly on an international network outside the regulatory scope of the domestic regulators. Second, recent trend suggests that the approach of national regulators is not unified.[164] While some countries have welcomed the financial innovation and reviewed their laws to regulate its activities (for example, United States) , others have either banned its operations or remain indifferent(for example, China).[165] The latter group's action is essentially premised on their believe that the VCs and exchanges are incapable of developing into a significant means of exchange over

[160] Masayasu Kanno, "Assessing Systemic Risk using Interbank Exposures in the Global Banking System" (2015) 20 Journal of Financial Stability 105; Ouarda Merrouche and Erlend Nier, "Payment Systems, Inside Money and Financial Intermediation" (2012) 21(3) Journal of Financial Intermediation 359.

[161] Alexander *et al*, *Global Governance of Financial Systems* (n 36) 23–6; *Interdependencies of Payment and Settlement Systems* (n 75).

[162] For a concise reading on present approach by states to the regulation of Virtual currencies, see Hughes and Middlebrook (n 6) 507–12.

[163] Defined Regulatory Arbitrage as "those financial transactions designed specifically to reduce costs or capture profit opportunities created by different regulations or laws": see Frank Partnoy, "Financial Derivatives and the Costs of Regulatory Arbitrage" (1997) 22 Journal of Corporate Law 211, 227, cited in Hughes and Middlebrook (n 6) 500. For further readings on arbitrage, see also, Annelise Riles, "Managing Regulatory Arbitrage: A Conflict of Laws Approach" (2014) 47 Cornell Law Review 63; Benjamin, *Financial Law* (n 94) ch 23.

[164] Hughes and Middlebrook (n 6) 507–12.

[165] *ibid*.

and above traditional payment systems.[166] Cross-pollination of risks between VC exchanges and TFIs connected to it make the likelihood of a cross-system risk more likely than ever if the financial consolidation considered becomes a reality. This draws from the same argument as the indirect interdependence of payment systems. For this reason, I argue that a domestic regulator-central bank-cannot efficiently manage the risk potentials stemming from the global network of VC exchanges operating independently or in consolidation with multinational TFIs.

In the next Part, I will consider the desirability of international bodies acting as regulators of VC exchanges and therefore managing the risks emanating from its operations as a global concern.

C. Should VC payment systems be regulated by international regulatory bodies?

Perhaps I should state presently that the term 'international regulatory bodies' is used in this Part to collectively refer to international organisations, institutions and agencies including: The Group of 20 (G20), Financial Stability Board (FSB), Basel Committee, Organisation for Economic Cooperation and Development (OECD), Bank of International Settlement (BIS), IOSCO, Word bank and International Monetary Fund (IMF). While their evolution and scope of application is admittedly invaluable, it is beyond the scope of this present paper.[167]Suffice to mention that their memberships account for more than 90% of central banks and key financial officers of both emerging and advanced economies.[168] My aim is to briefly examine why these bodies (including institutions and agencies which make up the international financial regulatory architecture) should be considered as the regulators of VC exchanges as a potential stakeholder in the interdependent global payment system.

In arguing the desirability of international bodies over domestic regulators, my primary focus will be on the need for international coordination and cooperation. This is because, empirical evidence from the recent financial crisis suggests that divergent approaches by states towards managing potential disruptions to global stability resulted in regulatory arbitrage.[169] This arbitrage

[166] *ibid* 514.

[167] For broad readings on the evolution and functioning of international financial institutions and organisations, see Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11); Alexander *et al*, *Global Governance of Financial Systems* (n 36); James Calvin Baker, *The Bank for International Settlements: Evolution and Evaluation* (Quorum Books, 2002).

[168] Brummer, *Soft Law and the Global Financial System* (n 80) 68.

[169] Christopher Arup, "The Global Financial Crisis: Learning from Regulatory and Governance Studies" (2010) 32(3) Law & Policy 363.

in turn fostered the disruption's navigation through several financial systems and the attendant systemic crisis failures.[170] A similar trend of divergent approaches are observable in the attempts at regulation of VC exchanges several financial systems.[171] To prevent the likelihood of another disruption navigating clandestinely through the interdependent payment network, there is need for cooperation and coordination among global regulators to manage potential disruptions.

International Coordination and Cooperation

As mentioned in Section II, global payment systems (and financial systems) are interdependent with the major drivers being technology, deregulation, and innovation.[172] The interdependence also means that potential risks and failures of each financial systems or GISIFIs could have adverse effects on other interconnected systems if not jointly managed.[173] These exposures to 'cross-system' risks make it incumbent on all financial systems and their regulators to cooperate (through 'dual supervision') in assessing potential risks to and jointly coordinating any risk management agenda.[174] The argument for the need of international coordination and cooperation for VC exchanges stem primarily from three reasons: (a) they operate through a web-based system globally and their risks are "a global problem that require global response";[175] (b) the attempt by national regulators are uncoordinated and overregulation could lead to regulatory arbitrage and transfer of systemic risks within the global payment network;[176] and (c) The authority of central banks are limited to their sovereign jurisdiction and inefficient for transnational supervision required for VC exchange activities.

The international bodies referred in this Section are mostly created in response to major crisis and disruptions that have threatened the global financial stability at one point or the other.[177] For example, the BIS (also termed the 'central bank's central bank') which manages the risks associated with international settlement and foreign exchange (with IMF, created during the Bretton Woods Agreement 1944) between payment systems was established in response to the

[170] *ibid*.
[171] Hughes and Middlebrook (n 6) 507–12.
[172] Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11) 10.
[173] *Interdependencies of Payment and Settlement Systems* (n 75).
[174] *ibid*.
[175] Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11) 16.
[176] Brummer (*ibid*) 227; Armour (*ibid*) ch 28.
[177] Brummer (*ibid*) 61; Alexander *et al*, *Global Governance of Financial Systems* (n 36) ch 2.

failure of German Bankhaus Herstatt earlier discussed.[178] Through the CPMI (formerly CPSS), the BIS has conducted several surveys on trends in payments and settlement systems to identify potential disruptions including recently on VCs.[179] It (like most international bodies) also has the general mandate of promoting the safety and efficiency of payment, clearing and settlement arrangements within and across financial systems.[180] Similarly, the G20 (formerly G7) comprising the largest economies globally was created to contain the systemic failures resulting from the 1997 Asian financial crisis which started Thailand.[181]

VC exchanges (like TFIs) perform core payment functions within multiple financial systems which may one day be crucial to stability of financial systems. The internationality and dominance of their operations may also be amplified if the integration with the TFI becomes a reality. If so, the internal complexities associated with the operations of exchanges and possible exploitation by TFIs will arguably make attempts by individual jurisdictions in regulating the institutions Herculean and inefficient.[182] Thus, they may become "too big to manage" (a subset of the notorious 'too big to fail').[183] The likely cross-pollination of risks from an important VC exchange to other institutions (and indirectly their financial systems) could result in potential disruptions to the global payment systems.[184] If this occurs, individual efforts of central banks will have little effect in containing the economic disruptions and spill-over except there is a communal approach between financial systems regulators globally. The communal approach canvassed here is one of the core features of international regulatory bodies.[185] The different bodies in the international financial architecture, while separate, complement one another through information-sharing, synchronised objectives, joint surveys,

[178] Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11) 75.

[179] Brummer (*ibid*) 85; Bank for International Settlements (Committee on Payments and Market Infrastructures), *Digital Currencies* (November 2015) <https://www.bis.org/cpmi/publ/d137.pdf> (accessed 9 August 2018).

[180] Bank for International Settlements (Committee on Payments and Market Infrastructures), *CPMI Charter* <https://www.bis.org/cpmi/charter.htm> (accessed 9 August 2018); *Payment Systems Worldwide* (n 15) 115; Armour, *Principles of Financial Regulation* (n 11) 621.

[181] For concise readings on the Asian crisis, see Roman Terrill, "The Promises and Perils of Globalization: The Asian Financial Crisis" (1999) 9(1) Transnational Law & Contemporary Problems 275.

[182] Fariborz Moshirian, "The Global Financial Crisis and the Evolution of Markets, Institutions and Regulation" (2011) 35(3) Journal of Banking and Finance 502.

[183] *ibid*.

[184] *Interdependencies of Payment and Settlement Systems* (n 75).

[185] Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11) 65–75.

implementation and compliance.[186] To facilitate the symmetry, they are classified broadly into agenda setters, standard setters, compliance monitors and sector setters.[187]

On the part of the participating economies, they function as implementers, peer reviewers and compliance monitors among themselves to ensure uniformity in compliance of the recommended standards and best practices to issues deliberated during sessions.[188] Compliance are ensured through binding laws by international institutions created by charter or statute (IMF, World Bank) and economic sanctions for breach.[189] However, save the two institutions mentioned, the predominant bodies in the international financial structures are established through international agreements, bylaws and declarations.[190]

This method has two significant implications: (a) Their existence and legitimacy are entirely dependent on the voluntary participation of member states; and (b) They have no powers to make binding laws but instead regulate through recommendations or published minimum standards which members voluntarily implement by reviewing the laws in their domestic financial systems (termed 'international soft laws').[191] The functioning of the international bodies considered above are crucial to the managing of potential disruptions that may affect the stability of financial systems (including payment systems). For VC exchanges, these bodies are through their global network able to gather information from several financial systems essential to the understanding of the dynamics of the innovative payment system.[192] This way, a uniformed regulatory approach could be devised through legislations and 'soft laws'. These will in turn be implemented

[186] Armour (*ibid*).

[187] Agenda setters are "organizations with broad and diverse memberships, including both regulatory and to, a lesser extent, political actors that are tasked with defining broad, strategic objectives for the international system'. Standard setters are lesser less political actors tasked with devising standards for ultimate adoption or implementation by national regulators and authorities. International monitors are 'organizations responsible for identifying whether (and to the degree to which) regulators comply with international financial law": see Brummer, *Soft Law and the Global Financial System* (n 80); Armour (*ibid*) 70, 74, 91; Alexander *et al*, *Global Governance of Financial Systems* (n 36) ch 2.

[188] Brummer (*ibid*) ch 2.; Alexander *et al* (*ibid*) ch 3.

[189] Brummer (*ibid*).

[190] *ibid* 63–4.

[191] *ibid*; "International soft laws refer to legal norms, principles, codes of conduct, and transactional rules of state practice that are recognized in their formal or informal multilateral agreements": see Alexander *et al*, *Global Governance of Financial Systems* (n 36)138.

[192] Brummer, *Soft Law and the Global Financial System* (n 80) 69.

globally by participating central banks with compliance monitored closely by both international agencies and peer review by member states.

However, as appealing as this strategy appears, it is not without criticisms that can limit its efficiency. First, since international regulatory bodies are comprised of both emerging and advanced financial systems, it raises a query on whether any approach adopted will suit either system.[193] This is especially because the intricacies of domestic financial systems are shaped significantly by the level of development of its financial markets, level of exposures of its consumers and political ideologies.[194] To have all these interests represented in the international forum results in significant clash of interest and political manoeuvrings.[195]

Second, a significant proportion of rules by international bodies are soft laws which are non-binding but subject to the voluntary implementation by member states.[196] Peer review and monitoring could be limited by states clinging to territorial sovereignty and economic protectionism which prevents a unified implementation of approach.[197] The challenges however do not, in practice, affect the effectiveness of the functioning of international financial bodies. States are in fact persuaded to cooperate amongst themselves because it guarantees better economic development to their financial systems.[198]

In conclusion, it is arguable that the limitations of the international regulatory bodies stated above could be easily circumvented through an increased autonomy to states in the implementation of the recommendations. This could be achieved by allowing member states to implement a slightly altered approach to the general recommendations by international bodies (termed 'subsidiarity').[199] This is because regulators of domestic systems—central banks—have better knowledge of the peculiarities of their financial systems and how best to manage potential disruptions.[200] This first-hand knowledge is not available to international bodies since they rely predominantly on information shared by members for conducting surveys into global concerns.[201] Most of the information may be doctored by central banks to protect the integrity of their payment systems and state sovereignty.[202]

[193] *ibid* 65–6.
[194] *ibid*.
[195] *ibid*.
[196] Alexander *et al*, *Global Governance of Financial Systems* (n 36) ch 4.
[197] Alexander *et al*, *Global Governance of Financial Systems* (n 36) ch 4.
[198] *ibid*.
[199] Brummer, *Soft Law and the Global Financial System* (n 80); Armour, *Principles of Financial Regulation* (n 11) ch 3.
[200] Brummer (*ibid*) 69.
[201] *ibid* 61–3.
[202] *ibid* ch 5.

Subsidiarity could admittedly create a situation where states abuse the privileges for political motives.[203]

But, this could be curtailed by subjecting any modification to strict compliance to the universal objective upon which such recommendations are made.[204] Furthermore, joint monitoring effort by regulatory agencies and peer review could also help ensure uniform implementation of the recommended best practices.[205] It remains to be seen whether the international community will move past the present dismissive attitude towards potential disruptions of VC exchanges to the global payment systems and take steps to manage the exposures while harnessing the economic benefits it offers.

V. Conclusion

This paper has identified what payment systems are, the categories and core functions they perform through the functioning of traditional financial institutions. The role of payment systems and financial institutions as key stakeholders in the global financial system stability was also discussed in the paper to exemplify the importance of their activities. I further examined the taxonomy of virtual currencies as payment systems, how their open distributed ledger network functions over p2p networks and importantly their functioning as payment systems through crypto-currency exchanges. Due to the role of payment systems in global stability, the paper also considered how the global payment systems work with focus on the role of multinational financial institutions as stakeholders of global payment system. The indirect interdependence of financial systems through the activities of financial institutions among other major drivers and its implications were identified. This was to establish a trend of how traditional financial institutions have survived as key stakeholders despite the recurrent innovations in payment systems.

As a foundation to the possible integration of virtual currency payment exchanges into this indirect relationship, the paper identified the major drivers of global interdependencies particularly technological innovation and financial consolidation between traditional payment systems and new entrants. The paper thereafter analysed the transactional advantages of using blockchain in the facilitation of large value payments as a possible driver for the integration between virtual exchanges and traditional institutions to benefit from the economies of scale and scope. To balance the scale, the attendant exposures of such integration, particularly the vulnerability of virtual exchanges to cyber-attacks and the systemic

[203] *ibid*.
[204] *ibid* ch 3.
[205] *ibid*.

risks to connected institutions was used as the baseline for the likely implications of a financial consolidation on the interdependent global payment systems.

Systemic risks arising from interdependencies of financial systems breeds a new form of cross-system failures where financial systems are exposed to the risks of other financial systems through the activities of financial institutions operating within multiple jurisdictions. The traditional risks of payment systems- payment and settlement risks and the knock-on effect it could have on connected systems and institutions were exemplified by the failure of Mt. Gox within the virtual community and Bankhaus Herstatt in the real economy. While these risks may be eliminated through blockchain system, it leaves open the possibility of exposure by the traditional institutions to the operational risks of VC exchanges if not closely managed. The risks, while possibly having immediate effect on the connected financial institutions, has the potentials of transferring to the financial systems in which they are popular.

Consequent upon the identification of these risks, the paper considered the possibility adopting a regulator that has the financial architecture to manage the exposures likely to emanate from the integration of the VC exchanges into the global payment network. A first reference was made to the activities of central banks as the major regulators of financial services providers in national payment systems. While their efforts are laudable, the paper argued from facts and examples that the limitations to the scope of authority resulting from sovereignty and lack of interoperability of central banks make them inefficient as the adequate regulator of VC systems. This conclusion was premised on the fact that virtual exchanges operate on a global network which central banks are incapable of controlling. The divergent attempts by different financial system regulators also breeds regulatory arbitrage that may affect trade balances and foster activities that might generate risks within regulatory gaps.

To ameliorate the challenges of regulating traditional institutions and multinational VC exchanges, the paper argued for the adoption of international bodies as the major regulators of VC exchanges. The core of the argument was anchored on the need for cooperation and coordination among financial systems to manage the global concerns arising from the operations of VC exchanges and by extension the payment instrument. The limitations of international bodies, particularly relating to the predominance of non-binding soft laws and concerns of political manoeuvrings between emerging and advanced economies during sessions was identified. As a possible solution to this challenge, the paper suggested a form

of subsidiarity where states are giving autonomy but subjected to the overriding good-faith towards the objectives of the international soft laws.

I argue that the views expressed, if adopted, will foster an enabling environment to usher in an era of global payment systems transformation through financial inclusion, transparency and efficiency. Risks arising from financial activities are not abnormal, the only abnormality arises from not paying attention to them until they result in catastrophic situations. The reactive approach to regulation is grossly inadequate to match the meteoric evolution of financial systems and activities. The integration of VC exchanges and traditional financial institutions will no doubt benefit the global payment systems and the broader financial system whose financial assets they facilitate if financial innovation is fostered and the risks jointly managed.

Cambridge Law Review (2019) Vol IV, Issue ii, 68–82

Old Is Sometimes Better: The Case for Using Existing Law to Face the Challenges of the Digital Age

Riccardo de Caria*

I. From "zero to one"? the "complicated relationship" between law and technology

Traditionally, all revolutions in history sooner or later had to come to terms with lawyers, and with the law in general:[1] the English and American revolutions were based on law and the rule of law, the French revolution run over existing law and rebuilt a new one, but certainly did not dispense with the lawyers, and even the Bolshevik revolution or Maoism established a new law alternative to the previous respective paradigms, but still felt the need to extensively use the law to pursue their goals. As for the revolutions understood not in the political sense, but in the technical-scientific one, at least in modern times, they have not led to

* Assistant Professor, University of Turin, Italy. Ph.D. (Turin), LL.M. (LSE).

[1] The link between law and revolutions has been the subject of several studies: a classic reference is to HJ Berman, *Law and Revolution: The Formation of the Western Legal Tradition* (CUP 1983); see also, among many: E Surrency, "The Lawyer and the Revolution" (1964) 8(2) The American Journal of Legal History 125; Hendrik Hartog (ed), *The Law in the American Revolution and the Revolution in the Law: A Collection of Review Essays on American Legal History* (New York University 1981); Michael P Fitzsimmons, *The Parisian Order of Barristers and the French Revolution* (Harvard University Press 1987); David A Bell, *Lawyers and Citizens: The Making of a Political Elite in Old Regime France* (OUP 1994); E Kolla, *Sovereignty, International Law, and the French Revolution* (CUP 2017); A Di Gregorio, "A 'new' State and a 'new' legal order: the Bolshevik Revolution and its legal legacy one hundred years after the 'October'" (2017) Diritto pubblico comparato ed europeo 993; M Loughlin, *Political Jurisprudence* (CUP 2017) ch 4, 'Burke on Law, Revolution, and Constitution', 63ff; Peter Charles Hoffer, Williamjames Hull Hoffer, *The Clamor of Lawyers. The American Revolution and Crisis in the Legal Profession* (Cornell University Press 2018).

the overcoming of the law, if anything to its transformation and adaptation: the industrial revolution certainly provoked considerable turmoil in the world of law, but the fundamental structure remained unaltered, and the lawyers maintained their prestige intact even in the very changed circumstances, effect of the scientific-technological upheavals. Therefore, no previous revolution actually engaged in "killing all the lawyers", not even metaphorically.[2]

Instead, the current digital revolution would seem to put the law and the role of lawyers in serious crisis for the first time: as is well known, we would be faced with a scenario in which law and code are assimilated and confused,[3] for which a complete rethinking and updating of the law would be necessary to keep up with the emerging disruptive technologies. In this context, even the role of lawyers would have to be completely thought over: tight between legal tech, artificial intelligence, machine learning, smart contracts, and their replacement with machines,[4] they would have no other choice but to reinvent themselves as computer scientists, relinquishing command of their discipline to the relentless advance of increasingly intelligent and powerful computers.

In other words, the typical feature of innovation in the digital world and in the field of new technologies, captured in the spot-on title of a wonderful book by Peter Thiel, *Zero to One*,[5] would have transferred to the law. That is to say that, contrary to what happened with previous revolutions, in this case even law should be, or perhaps already is subject to a break in continuity, or a complete break with

[2] As goes the famous line by the fictional character Dick the Butcher in William Shakespeare's Henri VI (Part 2): during a discussion among rebels led by Jack Cade on what they should do once they take the throne, the following conversation takes place: "*Cade*. Be brave, then; for your captain is brave, and vows reformation. There shall be in England seven halfpenny loaves sold for a penny: the three-hooped pot shall have ten hoops; and I will make it felony to drink small beer: all the realm shall be in common, and in Cheapside shall my palfrey go to grass. And, when I am king, – as king I will be, – *All*. God save your majesty! *Cade*. I thank you, good people: – there shall be no money; all shall eat and drink on my score; and I will apparel them all in one livery, that they may agree like brothers, and worship me their lord. *Dick*. The first thing we do, let's kill all the lawyers. *Cade* Nay, That I mean to do. […] *Cade* I thank you, good people: you shall eat and drink of my score, and go all in my livery; and we'll have no writing but the score and the tally, and there shall be no laws but such as come from my mouth".

[3] The reference is to Lawrence Lessig, *Code and Other Laws of Cyberspace* (Basic Books 1999); the dichotomy between law and code was picked up by many authors, including Primavera De Filippi and Aaaron Wright in *Blockchain and the Law: The Rule of Code* (Harvard University Press 2018).

[4] See for instance N Sahota, "Will A.I. Put Lawyers Out of Business?" *Forbes* (New Jersey 9 February 2019) <https://www.forbes.com/sites/cognitiveworld/2019/02/09/will-a-i-put-lawyers-out-of-business> (accessed 25 October 2019).

[5] Peter Thiel with Blake Masters, *Zero to One: Notes on Startups, or How to Build the Future* (Crown Business 2014).

the past, which would lead to a completely new paradigm, where progressively all the law is reduced to binary logic.[6]

In this paper, I would like to question this approach: I do not want at all to support a position of resistance to the new technologies, or yearn for a slowdown in the digital revolution, and in particular for the law to hinder the increasingly rapid innovation. This is in fact the position taken by some,[7] but, quite to the opposite, I believe instead that innovation, however rapid, should not be hindered, but rather favoured as much as possible in terms of public policy choices. Rather, I intend to focus on the question of whether the emerging technologies, which are making the scientific and economic world make a leap indeed from "zero to one", also actually require a law that is reborn from scratch, and the reinvention of its foundations, in order to keep up with innovation.

It has been argued that the law must be endowed with new categories in order to accommodate the novelties that the technological evolution is producing at an unceasing and increasingly accelerated rate.[8] I submit, however, that the law we have is perfectly equipped to regulate an economic and technological framework that is clearly evolving very rapidly: it is arguably not necessary to resort to new categories to keep up with this evolution, or, in any case, it would always be preferable to check with great caution whether the new legal problems that have arisen with the new technologies cannot be resolved in a completely adequate way with the legal categories of the analogue world. Only if this careful analysis is unsuccessful will the search for new categories be warranted, but this is by no means a foregone conclusion, it has to be proven, and the burden of proof lies with those who advocate the need to reinvent the law.

In the following paragraphs, I will therefore argue that the one between law and technology can be a happy marriage, and not necessarily a divorce, where each party go their own way. To this end, however, in order for this "complicated relationship[9]" to work, as in any self-respecting marriage, some essential ingredients will be needed, which I will go on to analyse in the next paragraphs: something old (Section II), something new (Section III), something borrowed (Section IV), something blue (Section V). The final paragraph (Section VI) offers

6 See the reflections by A Lo Giudice, "The Concept of Law in Postnational Perspective", in LH Urscheler and SP Donlan (eds), *Concepts of Law. Comparative, Jurisprudential, and Social Science Perspectives* (Routledge 2016) 209.

7 See for instance Fritjof Capra and Ugo Mattei, *The Ecology of Law: Toward a Legal System in Tune with Nature and Community* (Berrett-Koehler 2015).

8 At least to a certain extent, it is the position taken for example by E Biber and others, "Regulating Business Innovation as Policy Disruption: From the Model T to Airbnb" (2017) 70(5) Vanderbilt Law Review 1561.

9 As in one of the possible options to select for "relationship status" on a previous Facebook version ("it's complicated").

my conclusive remarks. In my work, I will make ample reference to the works of Bruno Leoni, an Italian philosopher of law who put forward some extremely acute and profound reflections which, although written a few decades ago, are very useful to respond to the legal challenges posed by the emergence on the scene of the new technologies.

II. Something old

As anticipated in the introductory paragraph, arguably the current rise of the digital technologies is only to a certain extent "the start of something new".[10] Let me make a few examples.

The sharing economy undoubtedly raises questions the legal notions of ownership and property:[11] how adequate is the traditional property law to cope with the changing relationships between people and things? We no longer own many of the goods we use, but we rent or lease them, or have them in other forms of temporary detention (think of cars, computers, smartphones),[12] or however goods of increasing economic importance see us as simply licensees, not owners (think of software, cloud storage space and what this entails for our files, etc.).[13]

As a consequence, to use the traditional metaphor of the 'bundle of sticks',[14] less and less sticks are left to the 'owner', and more and more remain in the hands of corporate powers and multinational conglomerates, who retain control over the goods they sell or lease or license.

To be sure, this is not in fact a completely new phenomenon, or in any case a phenomenon of the "zero to one" type: the issue of limited in time rights over things has been addressed by important authors for quite some time,[15] and even before the digital revolution showed all the disruptive potential that it has shown more recently. Therefore, the shift from property to other forms of relationship,

[10] This is just another reference to popular culture.

[11] See for instance S Kreiczer-Levy, "Consumption Property in the Sharing Economy" (2016) 43 Pepp L Rev 61.

[12] See C Cain Miller, "Is Owning Overrated? The Rental Economy Rises" *The New York Times* (New York, 31 August 2014), p SR3 of the New York edition.

[13] According to Kroll's *Global Fraud & Risk Report. Forging New Paths in Times of Uncertainty* (10th annual edn 2017/18), p 10, in 2017: "For the first time in 10 years of reporting, information theft, loss, or attack was the most prevalent type of fraud experienced in the last year, cited by 29% of respondents, up 5 percentage points from 24% of respondents in the 2016 survey. This in turn was up 7 percentage points from 22% of respondents in the 2015 survey. Theft of physical assets or stock, long the most common type of fraud, was the second most frequently cited incident, suffered by 27% of respondents".

[14] On which see, for instance, DR Johnson, "Reflections on the Bundle of Rights" (2007) 32 Vt L Rev 247.

[15] *E.g.* R Caterina, *I diritti sulle cose limitati nel tempo* (Giuffrè 2000).

typically contractual and temporary, with things, appears if anything relevant in terms of comparative increase in space for the latter type of relationship compared to traditional property, but not something radically, intrinsically new.

Even considering one of the innovative technologies on which the attention of public opinion and specialists has increasingly been focusing, namely the distributed ledger technologies (DLTs) famously at the basis of the blockchain, Bitcoin and smart contracts, the conclusion is arguably the same. Various legal systems, in particular those apparently more attentive to the evolution of new technologies and more eager to direct this evolution towards desirable goals and within a precise legal framework, have introduced normative definitions of these new phenomena,[16] thus sanctioning, more or less consciously, the idea that they represent something irreducible to the old legal categories, that warrants the creation of new ones.

In truth, this does not appear to be the preferable approach:[17] the DLTs, the blockchain, the smart contracts can in fact be classified in the existing categories,[18] without the need to introduce new institutions whose relationship with the existing ones will inevitably be problematic and will take some time to be arranged in a satisfactory way, if it can actually be arranged at all. Smart contracts are in fact pieces of software, whose legal implications it is a very stimulating endeavour to investigate;[19] but also the blockchain and Bitcoin can arguably be traced back to

[16] A comprehensive and updated account is for instance the one provided by Chetcuti Cauchi Advocates' *Blockchain, Crypto & ICOs. A Legal Review of Leading Jurisdictions* <https://blockchain.chetcuticauchi.com/report/> (accessed 25 October 2019). For some reflections on the regulatory responses to the DLTs revolution, see R Herian, "Regulating Disruption: Blockchain, GDPR, and Questions of Data Sovereignty" (2018) 22(2) Journal of Internet Law 1, 1, 8–16.

[17] For the opposite view, see for instance K Werbach, "Trust, but Verify: Why the Blockchain Needs the Law" (2018) 33 Berkeley Tech LJ 487, and OY Marian, "Blockchain Havens and the Need for Their Internationally-Coordinated Regulation" (2019) 20 North Carolina Journal of Law and Technology. Quite surprisingly, Eu institutions found no need to rush to regulate the crypto market: see Bloomberg, "Europe Is in No Rush to Regulate Crypto Market, Officials Say" *Fortune* (New York, 8 September 2018) <http://fortune.com/2018/09/08/europe-cryptocurrency-regulation/> (accessed 25 October 2019).

[18] See R de Caria, "The Legal Meaning of Smart Contracts" (2018) 26(6) European Review of Private Law 731.

[19] See R de Caria, "The Definition(s) of Smart Contracts Between Law and Code" in M Cannarsa, LA Di Matteo and C Poncibò (eds), *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (CUP 2019) 19.

existing legal institutions, which may vary according to different legal systems,[20] but which in any case exist everywhere.

But the same holds true for all the main areas in which we see the proliferation of technological innovation and the consequent rapid increase in legal questions that this entails: e-commerce does not change the key assumptions of contract law, 3D printing is part of a well-established framework of intellectual property law, the same artificial intelligence, robotics and the Internet of Things produce new scenarios and generate new problems, including legal ones, but they can be addressed satisfactorily with the statutory instruments of contractual and non-contractual liability.

Ultimately, it is arguably necessary to recover the notion of "evolutionary law" [21] adopted by a not yet enough appreciated academic tradition, among scholars of both of law and economics, ranging from Carl Menger to Murray Rothbard and especially Bruno Leoni.[22] In particular, the latter, highlighting the merits of the customary production of law in both the Roman and the common law traditions, formulated the original and fascinating theory of "law as individual claim",[23] according to which law in its proper and noblest meaning is something that arises spontaneously from the free repetition of certain behaviours by many people over time.

In other words, what makes a prerogative, an individual claim, rise to the status of recognised right (and therefore of law) is the spontaneous recognition of the same by the affiliates of a given community: following this approach, we must conclude that the rules governing the digital age should also be derived from the customary principles stratified over the centuries, that could be profitably employed also in the digital environment.

Indeed, it is by definition to be ruled out the contention that the law should follow and adapt rapidly to new scientific and market developments. It is therefore an effort that should not be undertaken, because the law in the strictest, noblest

[20] About Italy see *e.g.* P Burlone and R de Caria, 'Bitcoin e le altre criptomonete. Inquadramento giuridico e fiscale' (*Istituto Bruno Leoni*, 1 April 2014, IBL Focus 234) <http://www.brunoleonimedia.it/public/Focus/IBL_Focus_234-De_Caria_Burlone.pdf> (accessed 25 October 2019).

[21] On which see A Gianturco Gulisano, 'Bruno Leoni tra positivismo e giusnaturalismo. Il diritto evolutivo' (2009) Foedus 87.

[22] For a discussion of the influence of the Austrian school of economics on legal theory, see M Litschka and K Grechenig, 'Law by human intent or evolution? Some remarks on the Austrian school of economics' role in the development of law and economics' (2010) 29(1) European Journal of Law and Economics 57.

[23] It is the title of a chapter of his masterpiece, *Freedom and the Law* (see the expanded 3rd edn, with a foreword by Arthur Kemp (first published 1961, Liberty Fund 1991). The following references are taken from this online edn, available at <https://oll.libertyfund.org/titles/920> accessed 25 October 2019).

and most genuine sense is something that is stratified and consolidated over time: it is consequential that, if the time elapsed is short or very short, by definition law cannot have been created, because there has not been enough time for individual claims to establish.

To this it is worth adding that it appears a lost battle for lawyers the one chasing the latest new technological developments: either the law arrives late, thus risking to appear blunt, or, on the contrary, its fervour in trying to keep pace with innovation, never letting it happen without previously regulating it, ends up stifling that same innovation.[24]

III. Something new

However, it would be extremely hasty to conclude, from the premises set out in the previous paragraph, that there is and cannot be anything new in the law with regard to the new disruptive technologies. While it is true that the law should not chase innovation, it is certainly also true that something new actually exists, on various fronts.

This appears undeniable: wealth is increasingly represented by intangible goods: not only have we witnessed the shift of wealth from real estate to movable property, but the movable goods that become increasingly valuable are generally of an immaterial nature[25]. Hence the need to protect the new digital wealth, starting from the personal data, and to reflect on their concentration (big data) and the modalities of their treatment and circulation.

From this point of view, one of the main topics of discussion in the world of law is the privacy of such data: from the way in which it is often addressed, it would seem to be an essentially new and unprecedented problem, but as is well known, the right to privacy was conceptualised as far back as in 1890,[26] only 12 years after Edison invented his light bulb, just to make an example. And this right fundamentally remains the same, in its essential tenets, even though the

[24] The problems arising for regulation when facing fast-developing new technologies was dealt with for instance by M Fenwick, WA Kaal, EPM Vermeulen, "Regulation Tomorrow: What Happens When Technology Is Faster than the Law" (2017) 6 American University Business Law Review 561, although with policy conclusions different than my own.

[25] See, already several years ago, Margaret M Blair and Steven MH Wallman (eds), *Unseen Wealth: Report of the Brookings Task Force on Intangibles* (Brookings Institution Press 2001).

[26] SD Warren and LD Brandeis, "The Right to Privacy" (1890) 4(5) *Harvard Law Review* 193.

technological scenario has evolved in a tremendous way from the age of the Edison light bulb.

Arguably, therefore, the regulatory rush, that has put Europe at the vanguard in reining in this field, does not appear to be fully justified:[27] if the single piece of data is potentially subject to appropriation in the proper sense, then the traditional rules on (immaterial) property exist and can be applied; otherwise, it will be without protection, similarly to materials that do not have access to copyright protection.

To be sure, what makes the data special is that it is essentially much more useful to those who have to acquire it than to those who own it: the data is not only or so relevant in itself, but acquires particular relevance because of its combination with a vast amount of data from other people.[28] However, even in this case the problem is not new in itself: what the technological society changes from the past, if anything, is the ease with which such data can be found, but since forever, or at least since the capitalist mode of production has consolidated,[29] producers have had an interest in having as much information as possible about their customers. And such an information-gathering effort can arguably be accommodated in the existing legal framework without the need for new rules.

However, there are some areas in which we actually see the raise of new issues: in particular, artificial intelligence opens up scenarios that may conflict with the traditional notions of subjectivity, personality, responsibility (one can think of neural networks reproducing the functioning of the human brain, of machine learning, and the internet of things).[30]

In these cases, some new questions undoubtedly arise, which were not even conceivable a few years ago, and which are therefore worthy of further investigation. New problems certainly call for new answers, but the question becomes where

[27] For a study of the difficulties in complying with the new General Data Protection Regulation, see Sean Sirur, JRC Nurse and H Webb, "Are We There Yet?: Understanding the Challenges Faced in Complying with the General Data Protection Regulation (GDPR)" in *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security* (ACM 2018) 88; see also the critique to the GDPR by TZ Zarsky, "Incompatible: The GDPR in the Age of Big Data" (2017) 47 Seton Hall Law Review 995. For a defence of pre-GDPR EU privacy laws, see C Kuner and others, "Let's not Kill all the Privacy Laws (and Lawyers)" (2011) 1(4) International Data Privacy Law 209.

[28] See Shaira Thobani, *Diritti della personalità e contratto: dalle fattispecie più tradizionali al trattamento in massa dei dati personali* (Ledizioni 2018).

[29] To use the words of one of its paramount critics, Karl Marx.

[30] To a certain extent, this was already the contention made by D Friedman, "Does Technology Require New Law" (2001) 25 Harvard Journal of Law & Public Policy 71, 85: "If what we mean by 'new law' is 'new legal rules at the level of generality of the rules now used to decide cases', it is clear that new technologies will at least sometimes require new laws".

these answers should be found: do they require new law, or is the existing law sufficiently well equipped to accommodate these new questions?

In the next paragraph, I will argue why it is appropriate to refer, as often as possible, to existing law. Here, I would just like to add a comment about the fact that we can usefully distinguish between hard law and soft law when approaching this matter.[31] In fact, admittedly, what could be beneficial in order to effectively deal with the digital age is not a new wave of hard law, but at most a new set of soft law, simply helping to put what is new into context, to define it, and to make sense of it in perspective.

From this point of view, the white and green papers of the European institutions are something useful, as are the various policy tools available to regulators to clarify a matter, without intervening in an intrusive way. Soft law instruments are particularly to be welcomed when they anticipate future enforcement practices by public authorities, so as to dispel the possible uncertainty.

Also, the perspective of regulatory sandboxes appears to be a useful road to go down: as was effectively explained, "a regulatory sandbox is a framework set up by a financial sector regulator to allow small scale, live testing of innovations by private firms in a controlled environment (operating under a special exemption, allowance, or other limited, time-bound exception) under the regulator's supervision. The concept, which was developed in a time of rapid technological innovation in financial markets, is an attempt to address the frictions between regulators' desire

[31] For a general discussion, see GC Shaffer and MA Pollack, "Hard vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance" (2010) 94 Minnesota Law Review 706.

to encourage and enable innovation and the emphasis on regulation following the financial crisis of 2007–2008".[32]

This notion has emerged in the world of fintech, but could be usefully used in all areas of law that have to do with innovation: it would be a new way of dealing with novelties, without recourse to pre-emptive hard law.[33]

IV. Something borrowed

Ultimately, what the law can usefully do in order to deal with the new emerging technologies is to draw on the experiences of the past, and thus make use of the existing law, handed down to us precisely from the past.

What we should do, in particular, is to borrow the existing law and apply it to new cases: when a new technology is created, the preferred operation by both *policy-makers* and interpreters would be to consider the legal categories existing in private law, and make use of those that can be usefully applied to establish a framework for it. The tendency that we can observe, for instance in the field of DLTs,[34] seems rather to be to engage in a contest between different jurisdictions in order to be the first one to dictate new rules, often introducing new legal categories, in the hope of triggering a process of imitation of their solutions.

But if the legal comparison and the possible loan or transplant of rules from one system to another can undoubtedly be profitable in practice, in this case it would be preferable for the various systems to frame the novelty in their own categories. These categories, in fact, reflect at least in part a consolidation of legal 'claims'—to use Leoni's notion—to which over time the community of reference has recognised legitimacy and protection. It appears a wise choice in terms of policy to borrow such claims in order to frame the new ones that arise with the emergence of disruptive technologies.

This approach could face an objection if we consider, for example, a field such as the sharing economy, and in particular the case of Uber: borrowing the rules of other sectors could lead, for example, to believe that we have to apply the

[32] I Jenik and K Lauer, "Regulatory Sandboxes and Financial Inclusion" CGAP Working Paper, October 2017, <https://www.cgap.org/sites/default/files/Working-Paper-Regulatory-Sandboxes-Oct-2017.pdf> (accessed 25 October 2019).

[33] See various reflections in Financial Conduct Authority, Regulatory sandbox, November 2015, <http://www.ifashops.com/wp-content/uploads/2015/11/regulatory-sandbox.pdf> (accessed 25 October 2019); DW Arner and others, "FinTech and RegTech in a Nutshell, and the Future in a Sandbox" (2017) CFA Institute Research Foundation, 16ff; DA. Zetzsche *et al*, "Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation" (2017) 23 Fordham Journal of Corporate & Financial Law 31.

[34] See for instance R Girasa, *Regulation of Cryptocurrencies and Blockchain Technologies. National and International Perspectives* (Palgrave Macmillan 2018), especially ch 8, *International Regulation*, 199ff.

rules of labour law to the relationship between this company and its drivers, as is in fact invoked by the drivers themselves in different disputes in various jurisdictions.[35] Moreover, quite surprisingly, it was this very company that called, at policy level, for the introduction of new rules, so as to have a clear and defined regulatory framework within which to operate,[36] as if the existing law did not already contain rules that legitimised its action, if not at the level of ordinary legislation, at least at the constitutional level.[37]

Moreover, Uber repeated this strategy also in court, when it insisted on arguing that it carried out an activity that could not be framed as public passenger transport, which required instead specific treatment by the law (possibly, precisely, ad hoc treatment, and in any case new). This strategy has notoriously turned out to be a losing one before the Court of Justice of the Eu, which has instead found an identity between the activity carried out by Uber and the public transport activity carried out by existing operators, thus confirming that the existing law is able to make room for many of the innovations that are emerging, without the need for legislative interventions ex *novo*.

It is my contention that a more incisive strategy, even though admittedly not necessarily poised for success, would have been one based on acknowledging the facts, namely that the Uber directly competes with public transport operators, while at the same time claiming that it is (already now) allowed do so on the basis of the unalienable freedom of economic initiative, as well as of an immanent general principle of competition, which trumps any contrasting lower-ranking legislation.[38]

Consequently, borrowing the existing legislation appears to be the preferable choice, both in terms of general public policy, and for the disruptive companies involved themselves: deferring to the legislator for new legislation means putting oneself in the hands of the latter, with the risk that it will never decide, and in

[35] For a useful overview, see K Vizjak, "Uber – An Overview of the International Case Law" in Janja Hojnik (ed), *Sharing Economy in Europe: Opportunities and Challenges* (Zavod 14, 2018) 89ff.

[36] See for instance this report by C Zillman, "This Uber Exec Says the Startup 'Wants to Be Regulated" *Fortune* (New York, 14 June 2016) <http://fortune.com/2016/06/14/this-uber-exec-says-the-startup-wants-to-be-regulated/> (accessed 25 October 2019). For a quite opposite reading of Uber's regulatory policies, see R Collier, V Dubal and C Carter, "Disrupting Regulation, Regulating Disruption: The Politics of Uber in the United States" (2018) 16(4) *Perspectives on Politics* 919.

[37] I made this point, with regard to the Italian case (but the argument is valid in any other Western jurisdiction), in R de Caria, 'Profili di illegittimità nella disciplina italiana del trasporto pubblico non di linea' (2015) OPAL – Osservatorio per le autonomie locali N. 7.

[38] See M Delsignore, *Il contingentamento dell'iniziativa economica privata. Il caso non unico delle farmacie aperte al pubblico* (Giuffrè 2011).

any case that opportunities will be created for a very strong lobbying war between opposing positions, which inevitably ends up corrupting the law.[39]

V. Something blue

In the traditional matrimonial iconography, the blue item is meant to represent especially the value of purity, and fidelity between the bride and the groom. In our metaphor, the relationship between law and technology should be built in the spirit of fidelity by technological innovations to established law, that can thus preserve its purity.

As argued at the end of the previous paragraph, the value of purity of the law cannot give in to any need of the moment, however urgent it may appear. No matter how great the problems raised by the new technologies may appear, we must remain faithful to the idea that the law can evolve gradually, without shake-ups.

As Leoni put it, "Clearly 'legal' demands on one hand, and clearly 'illegal' demands on the other are located at the opposite ends of a spectrum comprising all demands that people may make in any given society at any given time. One should not forget, however, the huge intermediate sector of less definable 'quasi-legal' or 'quasi-illegal' demands whose probabilities of being satisfied are lower than those of clearly 'legal' demands, but still higher than those of clearly 'illegal' ones. The position of many, if not all, demands in the spectrum may change and is actually changing in any society at any given time. This process, to use Justinian's famous words, '*semper in infinitum decurrit*' (is always continuing), and we could not grasp it without introducing the time dimension. New demands may appear while old ones fade away, and present demands may change their position in the spectrum. The whole process may be therefore described as a continuous change of the respective probabilities that all demands have to be satisfied in a given society at any given time".[40]

Our faith should therefore be placed with the steady, and slow process of creation of the law, rather than with legislation: "This is certainly due, among other things, to the conventional faith of our time in the virtues of 'representative'

[39] In Leoni's words: "legislation is traced back, more or less implicitly, to the unconditioned will of a sovereign, whoever he may be. The very idea of legislation encourages the hopes of all those who imagine that legislation, as a result of the unconditioned will of some people, will be able to reach ends that could never be reached by ordinary procedures adopted by ordinary men; that is, by judges and lawyers. The usual phrase by the man in the street today, 'There ought to be a law' for this or for that, is the naive expression of that faith in legislation. While the processes conducive to lawyers'-law and judge-made-law appear as conditioned ways of producing law, the legislative process appears, or tends to appear, to be unconditioned and a pure matter of will' (Leoni (n 23)).

[40] Leoni (n 23).

democracy, notwithstanding the fact that 'representation' appears to be a very dubious process even to those experts on politics who would not go so far as to say with Schumpeter that representative democracy today is a 'sham'. This faith may prevent one from recognising that the more numerous the people are whom one tries to 'represent' through the legislative process and the more numerous the matters in which one tries to represent them, the less the word 'representation' has a meaning referable to the actual will of actual people other than that of the persons named as their 'representatives'".[41]

Policy-wise, this also implies that the regulation of digital technologies should be better left within the field of private law: the current expansion of the public law domain[42] is a move to reject, because it arguably leads to the interference of external interests, that are borne by somebody who is not involved in a transaction, in the free interplay of contractual wills by the parties to that transaction: "Dean Roscoe Pound pointed out in an essay cited by Professor Hayek that contemporary tendencies in the exposition of public law subordinate the interests 'of the individual to those of the public official' by allowing the latter 'to identify one side of the controversy with the public interest and so give it a great value and ignore the others.' This applies more or less to all kinds of administrative laws, whether they are administered by independent courts or not".[43]

In other words, injecting public law considerations into contractual relationships that should remain private is severely in danger of altering those relationships. As Leoni himself put it by making the example of marriage, particularly in tune with my own metaphor here, "That the legislators, at least in the West, still refrain from interfering in such fields of individual activity as speaking or choosing one's marriage partner or wearing a particular style of clothing or traveling usually conceals the raw fact that they actually do have the power to interfere in every one of these fields".[44] Admittedly, legislators have in fact the power to interfere and regulate the situations emerging from the new technological

[41] *ibid*.

[42] See the reflections on this topic by M Ruffert, "Public Law and the Economy: A Comparative View from the German Perspective" (2013) 11(4) International Journal of Constitutional Law 925.

[43] Leoni (n 23), citing FA Hayek, *The Political Ideal of the Rule of Law* (Cairo: Fiftieth Anniversary Commemoration Lectures, National Bank of Egypt 1955), 57 (later taken up in *The Constitution of Liberty*).

[44] Leoni (n 23).

advancements, but should arguably refrain from doing so, no less than how they normally refrain from interfering in marital relationships.

VI. Law and disruptive technologies: a happy marriage?

Ultimately, on the one hand, it may seem useful for lawyers to invent new categories in order to be ahead of the curve and not to give the feeling of being left behind, but it must always be stressed that the absence of new rules does not imply any regulatory vacuum *tout-court*,[45] and in any case even the regulatory vacuum is not in itself to be rejected.

In all modern legal systems, the "presumption of liberty"[46] should apply, according to which all acts not expressly prohibited by the law must be deemed lawful. In a rather peculiar move, a few years ago Italy decided to sanction this principle (even more oddly, only by way of ordinary law),[47] but this principle, even if implicit, is immanent to all democratic legal systems based on freedom and the rule of law. But if this is true, to go back to some of my examples, Uber does not—or at least should not—need new pre-emptive rules in order to know how to operate, or even to see itself legitimated to operate; Bitcoin of blockchain do not (or should not) need a normative definition in order to be able to be legitimately employed; artificial intelligence does not require a completely new legal paradigm and framework in order to be regulated; and so on.

I therefore submit that the law should not run after the latest technological advancement, but rather take a step back and effectively apply consolidated principles from the past even to today's fanciest technological gadgets In terms of policy, a strong self-restraint by legislators and regulators worldwide is in order: old-time customary law can arguably be much more effective in regulating the digital phenomena, than any hard-law attempt at reigning in an area that is irreconcilable with the uncertainties brought about by the tantrums of written law.

The use of the so-called new *Lex Mercatoria*[48] and customary law to tackle such hot topics as artificial intelligence, or internet of things, or robotics, might appear far-fetched. But what is stratified over time is what is really meant to last, overcoming the fashions of the moment. The most preferable approach seems

[45] In the field of DLTs, this view was shared for instance by IM Barsan, "Legal Challenges of Initial Coin Offerings (ICO)" (2017) Revue Trimestrielle de Droit Financier 54, 56, although the author reaches quite different policy conclusions than the ones advocated for here.

[46] On which see, among many, RE Barnett, *Restoring the Lost Constitution: The Presumption of Liberty* (Princeton University Press 2003).

[47] Art 3 of Decree-law 13 August 2011, No. 138, converted with amendments into law 14 September 2011, No 148

[48] For a comprehensive account, see, among many, A Stone Sweet, "The new *Lex Mercatoria* and transnational governance" (2006) 13(5) Journal of European Public Policy 627.

therefore to rely on the general "presumption of liberty", to admit that any new activity, however unusual and disruptive it may be, is lawful, unless it is expressly prohibited by an existing rule (but then by definition it will not really be so new), and to adopt a wait-and-see approach, possibly with the wide-ranging use, beyond the financial field, of the regulatory sandboxes.[49]

After all, the fact that certain areas of the law currently are comparatively less relevant than others, in terms of their economic relevance, does not change the nature of those areas of the law, let alone of the law in general itself: movable property may have become more valuable than immovable property, but legal categories remain the same that have been established over the centuries.

The relationship between technology and law can therefore be a happy and lasting marriage, provided that everyone maintains their own personality, and the bursting strength of the new technologies does not force the law to passively yield to the needs of the "partner": in order to avoid being cannibalised, the law must follow its own path, without compromising or losing its nature to chase the technology. Then the latter will find in the law strong, consolidated and useful categories, not destined to change within a short time-frame: only these can be solid bases for their relationship to be lasting and happy.

[49] We should indeed never underestimate the adverse effect on innovation – and thus economic cost – that a vast application of the precautionary principle can have, for instance as applied to tort law: see G Parchomovsky and A Stein, "Torts and Innovation" (2008) 107 Michigan Law Review 285.

Cambridge Law Review (2019) Vol IV, Issue ii, 83–104

The Enabler Theory and Atrocity Crimes

Shahram Dana*

I. Introduction

International criminal law (ICL) practitioners and scholars have observed that individuals convicted of atrocity crimes of similar gravity are sentenced to punishments of vastly different severity.[1] This raises questions whether "gravity" is indeed the primary consideration and differential factor in determining the quantum of punishment for atrocity crimes.[2] Is gravity of the offence operating as a meaningful differential principle in punishing atrocities? Is there an explanation that might reasonably justify substantially different sentences for persons convicted of crimes of similar gravity? Moreover, from a systemic perspective, has the notion of "gravity" been overplayed as a differential criterion for the purpose of punishing atrocities? And, has this come at the expense of developing sentencing criteria *sui*

* Senior Lecturer at Griffith Law School and Director, Law Futures Scholarly Roundtable Series at the Law Futures Centre, Griffith University, Australia. Former United Nations legal officer in the Office of the Prosecutor at the International Criminal Tribunal for the former Yugoslavia, Commissioner on the Torture Inquiry and Relief Commission, and law professor in Chicago, USA.
I am grateful to Nancy Combs, Mark Drumbl, and Robert Sloane for their comments on earlier drafts.

1 See Pascale Chifflet and Gideon Boas, "Sentencing Coherence in International Criminal Law: The Cases of Biljana Plavsic and Miroslav Bralo" (2012) 23 Criminal Law Forum 135, 147, 154; Jean Galbraith, "The Good Deeds of International Criminal Defendants" (2012) 25 Leiden J Int'l L 799, 800; Ines Monica Weinberg de Roca, "Sentencing and Incarceration in the Ad Hoc Tribunals" (2008) 44 Stanford J Int'l L 1, 6–12; Mark A Drumbl, *Atrocity, Punishment and International Law* (CUP 2007) 11, 15, 56–7; Mark B Harmon and Fergal Gaynor, "Ordinary Sentences for Extraordinary Crimes" (2007) 5 J Int'l Crim Justice 683, 684, 710; Andrew Keller, "Punishment for Violations of International Criminal Law" (2001) 12 Indiana International & Comparative Law Review 53, 65–6.

2 Robert Sloane, "Sentencing for the 'Crime of Crimes'" (2007) 5 J Int'l Crim Justice 713, 734 (rejecting the idea that gravity of the offence is a principle determinant of sentencing for atrocity crimes).

generis to atrocity crimes? This article explores these questions and responds with an original claim: the enabler theory.

In order to answer these questions, the research sought to identify individuals convicted of the same crimes. This methodology allowed the analysis to hold constant gravity of the crime as a sentencing factor permitting an evaluation of the influence of other factors on the quantum of punishment. It also provided a pathway to evaluate whether gravity genuinely operates, with consistency, as the *primary* factor influencing sentencing outcomes. This approach made significant the jurisprudence of the Special Court for Sierra Leone (SCLS) due to the prosecutor's distinctive approach to charging atrocity crimes. The SCSL Prosecutor conducted a single trial for each warring party in the armed conflict and prosecuted multiple co-perpetrators of varying rank within the same armed group under a single indictment with the same underlying facts and violations of law. Thus, in each trial, the same criminal charges were laid against all defendants, regardless of rank held within the group's hierarchy. This afforded a rare opportunity to analyse sentencing outcomes where perpetrators occupying different positions in the hierarchy of an organization are held criminally responsible for the same underlying crimes.

Furthermore, very few studies focus on the SCSL and its sentencing jurisprudence, in particular, has been largely ignored in academic literature.[3] By selecting the SCSL's jurisprudence as the subject of study, this article makes a significant and new contribution to ICL literature. This study underscores the importance of research analysing the rich jurisprudence of the SCSL, its significant contribution to the development of ICL, and its enduring relevance to the work of the International Criminal Court (ICC). The ICC has cited the SCSL sentencing decisions giving it continuing immediate relevance to ICL.[4] Additionally, among the *ad hoc* and hybrid tribunals, the SCSL enjoys the distinction of being the only international court to sentence a former Head of State. Thus, its jurisprudence includes the rare opportunity to examine the application of ICL sentencing law, norms, and principles to a head of state.

Drawing on the jurisprudence of the SCSL, this article offers an original claim regarding atrocity sentencing. I theorise that judges consider an atrocity criminal's role in enabling the context, that is conflict, in which atrocity crimes erupt when allocating an appropriate punishment. I call this the "enabler factor". Although ICL jurisprudence does not explicitly identify "enabling" as a specific sentencing factor, the notion is present in judicial narratives about the role of the

[3] But see Shahram Dana, "The Sentencing Legacy of the Special Court for Sierra Leone" (2014) 42 Georgia J Int'l & Comp L 615.

[4] *Prosecutor v Thomas Lubanga Dyilo,* ICC-01/04-01/06, Decision on Sentence pursuant to Article 76 of the Statute (10 July 2012) (hereinafter, "*Lubanga* Sentencing Judgment").

accused in the atrocities.[5] I argue that when ICL judges view an individual as an enabler, even if not explicitly articulated as a sentencing factor, they will punish that individual more severely than others that have committed similar crimes but were not enablers. ICL judgments in cases involving very powerful war criminals sometimes discuss the accused's role in enabling atrocities or the context of armed conflict that spawned atrocities crimes. This may appear in the court's judgment on guilty rather that in the sentencing judgment.[6]

The enabler theory closes the explanatory gap in sentencing outcomes left unexplained by the gravity narrative.[7] This article advances the enabler theory to explain perceived inconsistencies in the quantum of punishment between different trials, for example between the sentence of Charles Taylor and persons in the Civil Defense Force (CDF), a loosely organized fighting force loyal to President Kabbah and the ousted, democratically elected government of Sierra Leone. It also explains the distribution of punishment among co-perpetrators within a single trial. For example, in the RUF trial, the enabler factor explains why Issa Sesay received more than double the prison sentence of his co-defendant Augustine Gbao (52 years versus 25 years) even though both were convicted of the same crimes. This article also employs the enabler factor to explain the sizable differences between the punishment of individuals who fought against the government of Sierra Leone versus the comparatively light sentences of government supporters. Additionally, the enabler theory offers a pathway towards congruency between judicial narratives and actual sentencing allocations. The expressive function of atrocity trials and punishment is presently undermined by a singular focus on a narrative of extreme "gravity" that outpaces the actual quantum of punishment. The final sentences are underwhelming when compared to this hightened gravity rhetoric. It is also instructive to future sentencing determinations by the ICC and other international tribunals. In addition to closing the explanatory gap in sentence allocations, it

5 See, *e.g.*, Tadic Appeals Sentence, [55]–[58] (instructing trial judges to "consider the need for sentences to reflect the relative significance of the role of the Appellant in the broader context of the conflict"); *Prosecutor v Rukundo* (ICTR-2001-70 Trial Chamber), Judgment (27 February 2009) [605]; See also, *Prosecutor v Taylor*, Case No. SCSL-03-01-T, Judgment (18 May 2012) (hereinafter, "*Taylor* Trial Judgment") [5834]–[5835], [5842], and [6913]–[6915] (finding that "Taylor's acts and conduct had a substantial effect on the commission of the crimes because they: (i) enabled the RUF/AFRC's Operational Strategy; (ii) supported, sustained and enhanced the RUF/AFRC's capacity to implement its Operational Strategy.")

6 Compare *Taylor* Trial Judgment with *Prosecutor v Taylor*, Case No. SCSL-03-01-T, Sentencing Judgment (30 May 2012) (hereinafter, "*Taylor* Trial Sentencing Judgment").

7 Chifflet and Boas, *Sentencing Coherence in ICL* (n 1) 147, 156, and 158 (suggesting that gravity does not explain a large number of ICL sentencing outcomes).

also integrates sentencing outcomes with sentencing narratives and the goals of international prosecutions.

Section II launches straight into applying the enabler theory to explain the sentencing outcomes at the Special Court for Sierra Leone. It is examined in three distinct contexts. First, it is offered to explain the sentence of an individual atrocity perpetrator at the highest level of power and authority, such as a head of state. The enabler theory re-shapes the debate surround the punishment of Liberian President Charles Taylor, offering considerations that force rethinking of prevalent criticisms of his sentence. Second, through the prism of the enabler factor, I will illuminate the sentencing of three co-perpetrators of the same crimes in a way that "gravity" alone cannot explain. Third, the enable factor closes the explanatory gap between the sentences of perpetrators across the different trials and armed groups. Next, the article examines and responses to possible criticisms of the enabler theory. Section IV reflects on how existing ICL statutory provisions and sentencing factors can be interpreted to account for the enabler theory. The article concludes by examining some of the advantages of the enabler theory.

II. Enablers and mass atrocities in sierra leone

When heads of states or heads of armed groups enable mass conflict and criminality, an ominous environment for atrocities is created. The number of victims increases exponentially. Punishment in international criminal law must capture this extremely dangerous criminality. This can be done, for example, by giving weight the use of authority or power by head of a state or an armed group to enable atrocities, *i.e.*, to create or facilitate conditions that sustain atrocity criminality. This section applies the enabler factor to atrocities committed during a brutal war that fatally consumed the people of Sierra Leone for more than a decade.[8] In order to establish sufficient context, the section begins with a brief overview of the civil war in Sierra Leone.[9] This provides the factual context necessary for understanding the role of particular perpetrators and the application of the enabler factor. It then applies the enabler factor to the trials and punishments of: (1) Charles Taylor,

[8] For further reading on the conflict in Sierra Leone, see Ian Smillie, Lansana Gberie and Ralph Hazleton, *The Heart of the Matter: Sierra Leone, Diamonds & Human Security* (January 2000) <https://cryptome.org/kimberly/kimberly-016.pdf> (discussing the devastating nature of the conflict); Nicole Fritz and Alison Smith, "Current Apathy for Coming Anarchy: Building the Special Court for Sierra Leone" (2001) 25 Fordham Int'l LJ 391, 394 (*discussing* a campaign of terror against civilians that included abducting children, forced prostitution, and the amputation of limbs).

[9] This factual background borrows heavily from my earlier scholarship. See further, Shahram Dana, "The Sentencing Legacy of the Special Court for Sierra Leone" (2014) 42 Georgia J Int'l & Comp L 615, 619–22.

then sitting Head of State and President of Liberia; (2) three members of the Revolutionary United Front (RUF); and (3) two members of the Civil Defense Force (CDF).

In Sierra Leone's 1996 democratic elections, Alhaji Ahmad Tejan Kabbah, an ethnic Mandingo, was elected President of Sierra Leone, becoming his country's first Muslim Head of State.[10] Soon after the elections, armed conflict between the new government and the RUF resumed.[11] The RUF claimed that Kabbah's government was overrun with corruption, justifying an armed rebellion by the people.[12] From neighbouring Liberia, Charles Taylor long supported and enabled the RUF's war in Sierra Leone against the government of President Kabbah and past ruling regimes.[13] In December 1996, President Kabbah and RUF leader Foday Saybana Sankoh signed a peace agreement, the Abidjan Peace Accord, bringing a temporary halt to the atrocities and granting blanket amnesty to RUF fighters. However, peace did not last long. Within months, war consumed the country again. In March 1997, Sankoh was placed under house arrest in Nigeria for alleged weapons violations.[14] Although this gave Sierra Leone President Kabbah a small victory over the RUF, the ascendancy was short lived. A few months later, a group of senior military officers in the Sierra Leone Army (SLA) overthrew a weakened Kabbah in a *coup d'etat* and unlawfully seized power from the newly elected government with a brutal military assault on Freetown in which Liberian President Charles Taylor had "a heavy footprint" in planning, enabling,

[10] Charles C Jalloh, "Contributions of the Special Court for Sierra Leone on the Developments of International Law" (2007) 15 Afr J Int'l & Comp L 165, 169 ("Jalloh Contributions"). *Prosecutor v Sesay*, Trial Chamber Sentencing Judgment (8 April 2009) (hereinafter, "RUF Trial Sentencing Judgment") [146].

[11] Nsongurua J Udombana, "Globalization of Justice and The Special Court of Sierra Leone's War Crimes" (2003) 17 Emory Int'l Rev 55, 71 (*stating* that the atrocities were occasioned by the desire to control of the country's natural resources).

[12] See Babafemi Akinrinade, "International Humanitarian Law and the Conflict in Sierra Leone" (2001) 15 Notre Dame J L Ethics & Pub Pol'y 391, 392; Jalloh Contributions (n 10) 169; RUF Trial Sentencing Judgment (n 10) [146].

[13] Jamie O'Connell, "Here Interest Meets Humanity: How to End the War and Support Reconstruction in Liberia, and the Case for Modest American Leadership" (2004) 17 Harvard Hum Rts J 207, 213; Fritz and Smith (n 8) 394 (discussing how after the RUF entered Sierra Leone and controlled the Eastern region of the country, it implemented Charles Taylor's a campaign of terror by abducting children, forcing prostitution, and amputating limbs).

[14] *Prosecutor v Sesay, Kallon and Gbao*, Case No. SCSL-04-15-T, Trial Chamber Judgment (2 March 2009) (hereinafter, "RUF Trial Judgment") [19].

and overseeing.[15] Establishing themselves as the Armed Forces Revolutionary Council (AFRC), these mutinous officers installed one of their own, Johnny Paul Koroma, as Sierra Leone's new Head of State.

Although the AFRC and RUF joined forces, their union was an uneasy one. Together they fought against the CDF, which was led by Samuel Hinga Norman, an enormously popular figure and war hero among Sierra Leoneans.[16] With the intervention and support of the Economic Community of West African States Monitoring Group (ECOMOG), forces loyal to Kabbah, including the CDF, managed to regain control of Freetown and reinstate Kabbah as Sierra Leone's president. The civilian population of Sierra Leone, however, saw no respite from deliberate brutalization and horrific attacks against them. All parties to the conflict, including ECOMOG and Nigerian armed forces, continued to mercilessly attack, kill, and terrorize civilians. The retreating AFRC and RUF fighters looted and pillaged villages, killed or imprisoned civilians, and otherwise terrorized the population, including widespread mutilations and amputations. The hostilities and accompanying atrocities were intense, extreme, and prolonged.[17] After two more years of fighting, another peace agreement was signed and again RUF war criminals were granted full amnesty despite the horrible atrocities they committed. The 1999 Lome Peace Agreement, signed by President Kabbah and the RUF represented by Sankoh, not only gave Sankoh amnesty for atrocity crimes and pardoned his treason, but also installed him as Sierra Leone's Vice-President, and gave him control of the country's lucrative diamond mines.[18]

Two peace agreements and two full amnesties failed to deliver lasting or even short-term peace to the country, or respite to Sierra Leoneans from the horrors and hell they suffered. More hostilities followed and so too did graver atrocities. Throughout the war, Charles Taylor provided material support to the RUF/AFRC

15 *Taylor* Trial Sentencing Judgment (n 6) [76], [77], and [98]; *Prosecutor v Moinina Fofana, Allieu Kondewa*, Case No. SCSL-04-14-T, Trial Judgment (9 October 2007) (hereinafter, "CDF Trial Sentencing Judgment") [44]. See also Human Rights Watch, "Sierra Leone: Getting Away with Murder, Mutilation, and Rape" (1999) 11(July) Human Rts Watch 3(A) 12; James Rupert, "Diamond Hunters Fuel Africa's Brutal Wars" *Washington Post Foreign Service* (16 October 1999); Ian Stewart, "Rebels Set Freetown Ablaze, President Opens Talks" *Associated Press* (7 January 1999).

16 Antonio Cassese (ed), *The Oxford Companion to International Criminal Justice* (OUP 2009) 677.

17 See further, *The Heart of the Matter* (n 8); Fritz and Smith (n 8) 394 (2001) (discussing a campaign of terror against civilians that included abducting children, forced prostitution, and the amputation of limbs).

18 Tony Karon, "The Resistible Rise of Foday Sankoh" *Time Magazine* (12 May 2000) <http://www.time.com/time/arts/article/0,8599,45102,00.html> (accessed 3 July 2013); Obituaries, Foday Sankoh *The Telegraph* (31 July 2003) <http://www.telegraph.co.uk/news/obituaries/1437579/Foday-Sankoh.html> (accessed 18 March 2013).

armed groups that enabled them to continue the hostilities and atrocities, including supplying arms, weapons, munitions, and military personnel.[19] After the failures of the "peace with amnesty" strategy, movement towards accountability and justice gained traction, perhaps encouraged by the international tribunal model in response to the atrocities in Rwanda and Yugoslavia. In June 2000, President Kabbah requested that the United Nations Security Council establish a "special court for Sierra Leone" to prosecute RUF and AFRC leaders for planning and executing terrible atrocity crimes that brutalized and terrorized the people of Sierra Leone for more than 10 years.[20] The United Nations and Sierra Leone created a "special court" to prosecute persons bearing the "greatest responsibility" for the atrocities.[21]

The Special Court for Sierra Leone prosecuted and punished, *inter alia*, Charles Taylor, who was a sitting head of state when indicted; Alex Tamba Brima, Santigie Borbor Kanu, and Brima Bazzy Kamara, members of the AFRC Supreme Council and senior military commanders in the Sierra Leone Army that staged a successful *coup d'etat* that ousted the Sierra Leone government; Issa Sesay, senior military officer and commander of the RUF and later the combined AFRC/RUF forces in their insurrection against Sierra Leone; and Samuel Hinga Norman, founder and leader of the CDF who was appointed Deputy Minister of Defence by President Kabbah during the conflict.

In contextualising the enabler theory to the atrocities in Sierra Leone, it is helpful to first lay out the quantitative picture emerging from the sentencing practice of the SCSL. The Court has imposed nine sentences ranging from 15 years to 52 years with an average sentence of 36 years and median of 45 years. This picture changes dramatically if we examine separately the punishments of vanquished opponents of the Sierra Leone government and the punishments of the victorious pro-Sierra Leone forces. The average sentence for the vanquished opponents (*i.e.* Charles Taylor, the RUF and AFRC fighters) is 46 years; whereas the average sentence for supports of President Kabbah (*i.e.* the CDF defendants) is 17.5 years after appeal (the average sentence of these defendants at trial was 7 years), a mere fraction of the punishment met out to those that rebelled against the government. The CDF defendants also received the lowest individual sentences. Among the opposition groups, the AFRC, *i.e.*, the defecting military officers, were punished most severely with an average sentence of 48.3 years, comprising of

[19] *Taylor* Trial Judgment (n 5) [5834]–[5835], [5842], and [6913]–[6915].

[20] President of the Republic of Sierra Leone, Annex to the Letter dated 9 August, 2000, from the Permanent Representative of Sierra Leone to the United Nations addressed to the President of the Security Council, U.N. Doc S/2000/786 (10 August 2000).

[21] Jalloh, *Achieving Justice* (n 2) 398–404 (discussing the creation of the court).

individual sentences of 50 years for Brima and Kanu, and 45 years for Kamara.[22] The average punishment for the RUF defendants was 39 years.[23] Sesay received a prison sentenced of 52 years, the highest individual punishment rendered by the SCSL.[24] His RUF co-defendants Morris Kallon and Augustine Gbao received 40 and 25 years respectively.[25]

A. Enabler factor as applied to charles taylor

Several authors have criticised Taylor's fifty-year (50) prison sentence as excessive. Mark Drumbl, for example, argues that the judges were obsessed with Taylor's status as Head of State.[26] He concludes that this judicial obsession with accountability for a Head of State disproportionately increased Taylor's sentence.[27] Drumbl invites us to consider whether Head of State status should matter so much when it comes to punishing atrocities.[28] Likewise, Kevin Jon Heller argues that the mode of liability underlying Taylor's conviction, aiding and abetting, does not justify his sentence.[29] Heller's dissatisfaction stems largely from what he characterizes as poor reasoning and scant explanation for departing from basic principles that the judges themselves proffered as controlling the quantum of punishment.

Although subsequently overturned by the Appeals Chamber,[30] the trial judges took the position that aiding and abetting warrants a lesser punishment and furthermore proclaimed this to be a general principle of criminal law.[31] But then

22 See *Prosecutor v AFRC*, Case No. SCSL-04-16-A, Appeals Chambers Judgment (22 February 2008) (hereinafter, "AFRC Appeal Judgment") 105–6 (Sentencing Disposition).

23 See *Prosecutor v Sesay*, Case No. SCSL-04-15-A, Appeals Chambers Judgment (26 October 2009) (hereinafter, "RUF Appeal Judgment") 477–81 (Sentencing Disposition).

24 *ibid*.

25 *ibid*.

26 Mark Drumbl, "The Charles Taylor Sentence and Traditional International Law" *Opinio Juris Blog* (11 June 2012) <http://opiniojuris.org/2012/06/11/charles-taylor-sentencing-the-taylor-sentence-and-traditional-international-law> (accessed 14 April 2018) (hereafter "Drumbl, Punishing Heads of State (2012)").

27 *ibid*.

28 *ibid*; see also, Wayne Jordash and Scott Martin, "Due Process and Fair Trial Rights at the Special Court: How the Desire for Accountability Outweighed the Demands of Justice at the Special Court for Sierra Leone" (2010) 23 Leiden Journal of International Law 585 (claiming that the quest for accountability of perceived leaders has trumped certain fair trial rights).

29 Kevin Jon Heller, "The Taylor Sentencing Judgment: A Critical Analysis" (2013) 11 J Int'l Crim Justice 835 (hereafter "Heller, *Taylor Sentence* (2013)").

30 *Prosecutor v Taylor*, Case No. SCSL-03-01-A, Appeals Chamber Judgment (26 September 2013) (hereinafter, "*Taylor* Appeal Judgment") [666] and [670].

31 *Taylor* Sentencing Judgment (n 6) [100].

they immediately tossed this declared principle aside, and decide that they will instead consider the "unique circumstances" of Taylor's case when determining his punishment. The ruling comes across as an unjustified "go around" a purportedly basic principle of criminal law acknowledged by the Trial Chamber itself. Heller and others are rightly discontent. The convicted person, the victims, and the broader communities deserve a sounder justification.

There is some debate about whether the asserted rule—that aiding and abetting as a matter of law deserves less punishment—is a universally accepted general principle of criminal law.[32] Of course, factually, the actual contribution of an aider and abettor may be rather minor so as to warrant a lesser penalty. In such cases, a lesser penalty for aiding and abetting is the justifiable result of analysis and evaluation of the actual criminal conduct. It is not an automatic outcome. This is quite different from claiming at aiding and abetting as a matter of law merits a lesser penalty, regardless of the significance or decisiveness of the perpetrator's contribution. Facts must drive the analysis, especially in the context of atrocity crimes where the aiding and abetting is done by a head of state or head of an organised armed group. Yet, this critique does not resolve the problem here because, whether or not this is actually a general principle of criminal law, the Trial Chamber believed this rule to be a general principle and proceeded to sentence Taylor on that basis. So, assuming that this principle applies, is there a better explanation for the Trial Chamber's departure from it than the proffered "unique" circumstances of Charles Taylor case? If not, then perhaps Drumbl is right: "unique circumstances" might be simply a cover for a "fetish" to punishing a head of state.

One response might be that the critics are underestimating how seriously the Trial Chamber viewed Taylor's *planning* of crimes against humanity and war crimes. After all, Taylor was convicted of two modes of liability: planning as well as aiding and abetting.[33] Still, this explanation does not get us home because Taylor was convicted of "planning" only one of his crimes; for all others, he was convicted as an aider and abettor. An explanation with greater legs is that the

[32] The SCSL Appeal Chamber decisively concludes that it is not. It holds that the Trial Chamber's position is inconsistent with the SCSL's statute, rules, and jurisprudence. See, *Taylor* Appeal Judgment (n 30) [666] and [670]. The Appeal Chamber further holds that the Trial Chamber's ruling here is not supported by customary international law, nor a general principle of law, citing and discussing numerous jurisdictions including the Sierra Leone, the United States, Austria, Brazil, Costa Rica, Puerto Rico, France, and Italy: see [667]. The SCSL also persuasively demonstrates the error of arguments that rely on ICTY jurisprudence to claim that aiding and abetting as a mode of liability warrants lesser punishment: [666]–[669].

[33] *Taylor* Trial Judgment (n 5) [37B].

judges considered Taylor's contribution to the atrocity crimes to be a particularly serious form of complicity. They considered Taylor to be an enabler.[34]

A closer examination of the trial judgment reveals that the judges consider Taylor to have not merely aided and abetted in the crimes but in fact *enable* the atrocities.[35] He enabled the RUF/AFRC's operational strategy, which included the committed atrocity crimes, by supporting, sustaining and enhancing the RUF/AFRC's capacity to carry out these activities.[36] Taylor supplied arms and ammunitions to the RUF and AFRC that were "indispensable" in empowering them to launch attacks.[37] The RUF repeatedly encountered the problem of depletion of its arms and military supplies and time and again Taylor responded by directly supplying them with more weapons.[38] Without the shipment of weapon's from Charles Taylor, the RUF could not have sustained their attacks on civilians.[39] During the same time, the United Liberation Movement of Liberia for Democracy (ULIMO) was supposed to disarm and surrender its weapons to the UN. Instead, Taylor enabled them to sell or barter their weapons to the RUF, and further provided the RUF with the financial means needed to purchase these arms and ammunitions from the ULIMO.[40] Taylor also directly supplied the AFRC with arms.[41] The trial judges found that "Taylor's acts and conduct had a substantial effect on the commission of the crimes because [he] enabled the RUF/AFRC."[42]

While the Trial Judgment is explicit that Taylor "was critical in enabling" the RUF and AFRC,[43] the Sentencing Judgment does not explicitly identify "enabling" as a sentencing factor. Nevertheless, I argue the enabler factor influenced and increased Taylor's sentence. Implicit in the judges' sentencing narrative is their grave concern that Taylor enabled the armed conflict and ensuing atrocities. When the Trial Chamber analyses the "role of the accused", it considers Taylor's "sustained operational support;" "the steady flow of arms and ammunition;" and

[34] *E.g. Taylor* Trial Judgment (n 5) [6914] (finding that Taylor "was critical in enabling" the RUF and AFRC). See also, *Taylor* Appeals Judgment (n 30) [683].

[35] *E.g. Taylor* Trial Judgment (n 5) [5834], [5835], [5842], [6913]–[6915] (finding that "Taylor's acts and conduct had a substantial effect on the commission of the crimes because they: (i) enabled the RUF/AFRC's Operational Strategy; (ii) supported, sustained and enhanced the RUF/AFRC's capacity to implement its Operational Strategy.")

[36] *ibid*.

[37] *E.g. Taylor* Trial Judgment (n 5) [5834], [6914].

[38] *ibid* [5837], [6914].

[39] *ibid* [6913]. See also, *Taylor* Appeals Judgment (n 30) [683].

[40] *Taylor* Trial Judgment (n 5) [5835].

[41] *ibid* [5837].

[42] *ibid* [5834], [5835], [5842], [6913]–[6915]

[43] *E.g. ibid* [6914].

"the extended duration of the conflict… and crimes."[44] The judges further found that without Taylor's involvement "the crimes might have end earlier."[45] These finding concerning Taylor's role in the atrocities strongly indicate that the judges considered Taylor to be an enabler, and they further imply this by placing him in "a different category of offenders for the purpose of sentencing."[46] Clearly, the fact that Taylor was responsible for knowingly enabling a dire situation encouraging the commission of atrocities, some of which he planned himself and others in which he was complicit and made a decisive contribution, weighed heavily on the judges during sentencing deliberations.

The enabler factor explains Taylor's sentence by considering his role in enabling and maintaining a milieu or situation of atrocities. The judges found that Taylor enabled the commission of atrocity crimes in another state.[47] In the context of atrocity crimes and the goals of international criminal justice, a Head of State using his power and state resources to *enable*, as the SCSL describes it, atrocities and conflict in the territory of another country is the archetype criminality that ICL is most concerned with. It is arguably the quintessence of atrocity crimes. This wrongdoing must be accounted for in the punishment. The harm here goes beyond public international law concerns regarding state sovereignty. As a Head of State, Taylor's criminality decisively enabled the commission of crimes against humanity and other horrific atrocity crimes against the civilians of another state. Thus, when such extraterritorial criminality is committed by a person in control of a foreign state's armed forces or military resources, the nature of this criminality must be linked to the quantum of punishment. The *Taylor* Trial Chamber sought to capture this harm—the extraterritorial nature and effect of his wrongdoing—as an aggravating factor.[48] Such a conceptualisation is reasonable, but arguably could have been sharpened by more direct accounting of how Taylor used his power as a head of state to enable atrocities and connecting that dire harm to the resulting increase in punishment. The danger lies not simply in extraterritorial criminality, but in a very powerful actor enabling atrocities in a neighbouring country. The punishment is better explained by the enabler factor.

B. Enabler factor as applied to the RUF case

Comparing the sentences of Sesay, Kallon, and Gbao gives further traction to the influence of the enabler factor. Given that all three defendants were convicted of the same underlying crimes (acts of terrorism, mutilations and cutting

[44] *Taylor* Sentencing Judgment (n 6) [76].
[45] *ibid*.
[46] *ibid* [100].
[47] *ibid* [98].
[48] *ibid*.

off of limbs, rape, sexual slavery, murder, enslavement, pillage, forced marriage, and attacks on peacekeepers), this trial allows us to compare crimes of the same gravity. When we compare the sentence of Sesay, the man who had the power to effectuate the disarmament of the RUF, with Gbao, also a senior military commander reporting directly to Sesay, we learn that "gravity" does not function as a differential sentencing principle as judges claimed. For some crimes, Sesay was sentenced to three times the prison term that Gbao received, as shown in the table below.[49] For example, both were convicted of rape as a crime against humanity under Count 6 of the indictment.[50] For this offense, Sesay was sentenced to 45 years of imprisonment, whereas Gbao received only 15 years. For sexual slavery as a crime against humanity under Count 7, Sesay was sentenced to 45 years; Gbao got 15.[51] For pillaging as a war crime under Count 14, Sesay was sentenced to 20 years; Gbao got six.[52] For other inhumane acts as a crime against humanity under Count 7, Sesay was sentenced to 40 years; Gbao got 11.[53]

CRIME	**SENTENCE**
Rape (CAH)	Sesay – 45 years Gbao – 15 years
Sexual Slavery (CAH)	Sesay – 45 years Gbao – 15 years
Pillaging (WC)	Sesay – 20 years Gbao – 6 years
Terrorism (WC)	Sesay – 52 years Gbao – 25 years
Attacks against Peacekeepers (WC)	Sesay – 51 years Gbao – 25 years
Other inhumane acts (CAH)	Sesay – 40 years Gbao – 11 years

Furthermore, comparing their punishments for other crime reveals a similar pattern. Crimes for which Sesay received 52 years (terrorism as a war crime) and 51 years (attacks against peacekeepers), Gbao received only 25 years for the same crimes,[54] which were in fact charged in the indictment under the same count against both defendants (Counts 1 and 15).[55] Yet, Sesay received more than double

49 RUF Appeal Judgment (n 23) 477–81 (Sentencing Disposition).

50 *Prosecutor v Sesay, Kallon and Gbao*, Case No. SCSL-2004-15-PT, Corrected Amended Consolidated Indictment (2 August, 2006) (hereinafter, "RUF Indictment").

51 RUF Appeal Judgment (n 23) 477–81 (Sentencing Disposition).

52 *ibid.*

53 *ibid.*

54 RUF Appeal Judgment (n 23) 477–81 (Sentencing Disposition).

55 RUF Indictment (n 50).

the prison time that Gbao got for these crimes. Likewise, for extermination as a crime against humanity under Count 3, Sesay was sentenced to 33 years; Gbao got 15. Thus, the increase in the penalty for the same crime is more than 100%.

The Trial Chamber treated Sesay as the most influential and highest-ranking battlefield commander of RUF/AFRC; in other words, Sesay was the ultimate commander of the all rebel fighting forces.[56] Sesay's power to end the conflict was amply demonstrated when under his orders the entire rebel forces demobilized. Given Sesay's crucial role in sustaining the conflict and atrocities, and his 52-year sentence (the highest at the SCSL) compared to the 25 years Gbao received, the sentencing practice here indicates that the enabler factor will enhance punishment by more than 100%.[57] Thus, gravity of the offense, if understood as the seriousness of the harm, is not as controlling of quantum of punishment as the rhetoric of the SCSL and other international criminal courts suggests. In fact, gravity can hardly be considered a "litmus test" for a sentence when one defendant receives double or triple the sentence of his co-defendant when both were convicted of the very same crime—thus same gravity—alleging the same facts under the very same count in the indictment. For the same crimes, Sesay was sentenced to 27 years more than Gbao.[58] Attributing this increase of 27 years to the aggravating factor of Sesey's "position as a superior" does not convincingly account for the quantum of the increase. Gbao was also a senior commander.

The difference is also striking when we compare Sesay to Kallon. For the crime of acts of terrorism, Kallon got 39 years and Sesay received an additional 13 years, increasing his punishment to 52 years of imprisonment, a 33% increase in punishment.[59] Likewise, for the crime of attacking peacekeepers, Sesay's punishment was 51 years of imprisonment whereas Kallon received 40 years.[60] Thus, Sesay received 11 more years than Kallon for the same crime, making Sesay's prison sentence more than 25% longer than Kallon.

The sentences in RUF case demonstrates that even between high level perpetrators (Sesay, Kallon and Gbao), the punishment for the enabler among them increases quite substantially. The critical role of a very high-ranking accused in enabling and creating a milieu for systemised criminality is a weighty differential factor and can reasonably account for this sharp increase in penalty. The enabler factor is very significant in sentence allocations for atrocity crimes, even though it is often not clearly articulated in international judgments. So influential was the

[56] *ibid* [22] and [23].
[57] RUF Appeal Judgment (n 23) [1206] (finding that Sesay's highly influential role increased the gravity of the offences).
[58] RUF Appeal Judgment (n 23) 477–81 (Sentencing Disposition).
[59] *ibid*.
[60] *ibid*.

enabler factor at sentencing that, in the RUF case, it nullified the fact that Sesay's responsibility for the attacks on the peacekeepers was less than Kallon's culpability for the same crime. This suggests that the enabler factor is a more significant factor for sentencing than the accused's mode of liability. Kallon ordered the attacks on peacekeepers and he even personally attempted to kill an UNAMSIL officer.[61] These constitute particularly grave modes of liability. Kallon bore direct criminal responsibility for direct participation, according to the Trial Chamber.[62] Sesay's responsibility however was further removed and less culpable relatively speaking. The Trial Chamber found Sesay to be only indirectly responsible for the attack because he failed to punish individuals like Kallon who ordered and carried out the attacks.[63] Thus, Sesay's only culpability was by omission, compared to Kallon ordering and personally participating in the crime. Nevertheless, for the same crime, Sesay was sentenced to 11 more years of imprisonment than Kallon.

Although Sesay was not a head of state, he did direct and enable all RUF activities in Sierra Leone after Sankoh was imprisoned.[64] Thus, for the relevant time periods, he was the head of an organised armed group engaged in armed conflict against a state. I theorize that accounting for the enabler factor is implicitly what some international judges are doing in their determination of what constitutes a just and appropriate punishment, even if their sentencing judgments fail to explicitly articulate enabler as a sentencing factor and even despite their magniloquence about "gravity of the offense" as the dispositive criteria for atrocity sentencing. The enabler factor better explains the reason for the very substantial increase in Sesay's punishment compared to his RUF co-perpetrators. Responsibility for enabling atrocities is a significant differential factor in sentencing allocations for international crimes.

C. Enabler factor as applied to CDF case

The CDF defendants, Moinina Fofana and Allieu Kondewa, received the lowest sentences of any atrocity perpetrator convicted by the SCSL. The Trial Chamber sentenced Fofana to a meagre, six (6) years of imprisonment and Kondewa to eight (8) years for very heinous crimes including murder, cruel treatment, pillage, and using children in hostilities.[65] Thus, their crimes are quite

[61] RUF Trial Judgment (n 14).

[62] *ibid*.

[63] *ibid*.

[64] RUF Indictment (n 50) [23].

[65] CDF Trial Sentencing Judgment (n 15) 34–5 (Sentencing Disposition). The SCSL Appeals Chamber increased their sentences to 15 and 20 years respectively for Fofana and Kondewa. See *Prosecutor v Moinina Fofana, Allieu Kondewa*, Case No. SCSL-04-14-A, Appeals Chamber Judgment (28 May 2008) (hereinafter, "CDF Appeal Judgment") 189.

grave, like the crimes committed by other perpetrators convicted by the SCSL.[66] Yet, the CDF war criminals' punishment is drastically lower than the average sentence (48 years) for other trials at the SCSL. Does the "gravity of the offence" factor adequately explain sentences of six (6) and eight (8) years for murder, cruel treatment, pillaging, and using children in hostilities? Can the enabler factor better explain why the CDF defendants received substantially lower punishments than the defendants in the RUF and AFRC trials?

How did the judges arrive at these sentences? As all ICL sentencing judgments do, the trial judges in the CDF case narrate their sentences in terms of "gravity".[67] And, like all other ICL sentencing analysis, the judges do not follow a doctrinal approach to gravity. Instead, they list a number of factors relevant to the gravity assessment: "scale and brutality of the offenses committed, the role played by the Accused in their commission, the degree of suffering or impact of the crime on the immediate victim, as well as its effect on relatives of the victim, and the vulnerability and number of victims."[68] If gravity is the litmus test, then punishments of six and eight years of imprisonment seem manifestly low for perpetrators convicted of murder, forcing children to kill for them, and acts of rape resulting in death.[69] If gravity demands a more severe punishment, how do the judges justify this comparatively low sentence? Part of the explanation lies in the judges' sympathy for the reasons the CDF Kamajor fighters entered the armed conflict, namely in "defence of their communities… with the sole objective of… preventing the brutal killings" of their families and to "protect their lands and properties."[70] Another prominent narrative also influenced the sentence.[71] According to the judges, a significant justification for the low sentence was that these war criminals fought for a "legitimate cause" by preventing a *coup d'etat* against the elected government.[72] The CDF defendants helped "re-establish the rule of law" by defeating a rebellion.[73] Fighting for a legitimate cause merited mitigation of punishment, in the eyes of the judges, even if the means to achieve that end included atrocity crimes.[74] The sentencing narrative exudes a distinctive tone of redemption, rather than condemnation as might be expected

[66] For a complete discussion and analysis of all crimes and punishment of defendants before the SCSL *see* Shahram Dana, "The Sentencing Legacy of the Special Court for Sierra Leone" (2014) 42 Georgia J Int'l & Comp L 615.

[67] CDF Trial Sentencing Judgment (n 15) [33].

[68] *ibid*.

[69] *ibid* [47] and [52].

[70] *ibid* [84].

[71] *ibid* [91] and [94].

[72] *ibid* [82]–[94].

[73] *ibid* [87].

[74] CDF Trial Sentencing Judgment (n 15) [44] and [82]–[94].

in a judgment on criminality. The judges moralised that fighting to restore the "legitimate" government "atones" for their "grave and very serious" crimes.[75] Justifying reduction of punishment based on fighting for the good guys is deeply problematic, and the SCSL Appeals Chamber promptly overturned this ruling. Mitigating a sentence because the war criminal fought on the right side offends a core value of international humanitarian law, that all sides conduct hostilities in accordance with basic laws of humanity. To allow otherwise, destroys the *jus in bello, jus ad bellum* distinction. This justification for mitigation is liable to criticism of "victor's justice" and politicisation of international justice and runs counter to its underlying goals. The Trial Chamber's willingness to advance "legitimate cause" as grounds for mitigating punishment for atrocity crimes has been met with severe criticism.[76] The enabler factor, on the other hand, offers an explanation more congruent with the ethos of international criminal justice and the goals of international atrocity trials.

If gravity of the offence was indeed the primary factor influencing the trial sentence of the CDF defendants, then it would be reasonable to expect a higher sentence than six (6) and eight (8) years for murder, using children in hostilities, cruel treatment, and pillaging. Thus, the gravity factor does not explain these sentences. Could these sentences be better explained by the fact that Fofana and Kondewa were not enablers? Nowhere in the judgment of the CDF do you find the characterization of them as enablers as you do with, for example, Charles Taylor.[77] The Trial Chamber also noted "Fofana's commitment to and observance of the Lome Peace agreement", demonstrating commitment to non-conflict.[78] He also made substantial efforts to "ensur[e] that members of the CDF remained committed to the peace process."[79] The Trial Chamber commended Fofana's post-conflict efforts to foster the peace process.[80] These findings all indicate that the judges did not consider the CDF defendants to be enablers of the armed conflict that spawned the atrocities. The judges recognised that factors indicative of an accused's role as an enabler (or non-enabler) are significant in determining an appropriate sentence. They determined that Fofana and Kondewa merited a

75 *ibid* [82]–[94].

76 Human Rights Watch, Political Considerations in Sentencing Mitigation for Serious Violations of the Laws of War before International Criminal Tribunals (March 2008) <https://www.hrw.org/sites/default/files/related_material/Political%20Considerations%20in%20Sentence%20Mitigation%20for%20Serious%20Violations%20of%20the%20Laws%20of%20War%20before%20International%20Criminal%20Tribunals.pdf> (accessed 15 August 2019).

77 See also, *Taylor* Trial Judgment (n 5) [5834], [5835], [5842], [6913]–[6915].

78 CDF Trial Sentencing Judgment (n 15) [67].

79 *ibid*.

80 *ibid*.

sentence lower than what the gravity of the crime might otherwise demand. The fact that Fofana and Kondewa were not enablers, and in fact were viewed by the judges as persons committed to non-conflict, better explains their sentences than the gravity of their offences.

III. Possible critiques of the enabler factor

This article does not present the enabler factor as an exclusive explanation that completely responds to all that vexes the goal of justly punishing atrocities. I argue that it elucidates a real determinate of ICL sentencing that is not explicitly accounted for in the sentencing narratives, and thus the enabler theory brings greater clarity to punishment for atrocity crimes. It captures an atrocity perpetrator's responsibility for enabling and creating the context of armed conflict. This element is not yet accounted for in any explicit sentencing factor. Yet, atrocities generally do not occur absent the chaos and disorder of war or armed conflict. In addition to better explaining ICL sentences, the enabler factor also enriches the sentencing analysis, leading to greater coherence between judicial narratives and sentencing outcomes.

Nevertheless, I recognise some concerns and possible arguments against the enabler theory. First, some may argue that it is too narrow in its applicability. This critique points out that the enabler factor will only apply to a small percentage of war criminals and thus have limited utility. I argue that its limited applicability is actually a strength. The reality is that only a few war criminals have the capacity to be enablers of armed conflict.

Second, it may also be argued that the enabler factor unfairly backdoors responsibility for aggression, violating *nullum crimen sine lege*. This criticism misunderstands the scope of the enabler factor. It does not create an independent or separate grounds of individual criminal responsibility for the crime of aggression, or any other atrocity crime for that matter. The accused must still first be found guilty of an existing atrocity crime, thus respecting the principle of legality. Consideration as an enabler arises only after a finding of guilt for an existing crime for the purpose of determining an appropriate sentence.

Third, and related to second, is the criticism that the enabler factor appears to only capture the person or side that initiated the conflict. It is true that persons who initiate large scale armed violence are more likely to be captured by the enabler factor. I argue that such dire conduct is dangerous criminality that must and should be captured in the sentence. Moreover, nothing about the enabler factor necessarily limits it to initiators of armed conflict. It is quite possible that a non-initiator of the violence can be found to be an enabler. The analysis would be

driven by the facts on the ground, not labels. In the CDF case, the Kamajor were not viewed as enablers; the judges found that they did not enable the civil war in Sierra Leone but only engaged in the fighting in "defence of their communities" to protect their families from becoming victims of atrocities.[81] Based on these findings of fact, the Kamajor would not be considered enablers of Sierra Leone's civil war. Admittedly, the determination of who are enablers, and who are not, can be politicised. However, this does not undermine the enabler theory as a helpful tool for conceptualising the role of accused in sentencing narratives. It simply draws attention to the importance of objective application, which is true of any judicial determination. International judges are already called upon and indeed do make determinations on matters that are politicised and can impact domestic and international politics such as, for example, whether a war criminal contributed to national reconciliation.[82]

To be clear, the fact that a war criminal is not an "enabler" does not exonerate them from their crimes. They are still criminally responsible for their hand in the atrocities. A non-enabler can still be guilty of atrocity crimes and punished accordingly. The CDF defendants were found guilty of and punished for their crimes.

IV. How the enabler factor fits into ICL statutory provisions

How might international criminal courts incorporate the enabler factor into sentencing determinations? This question is beyond the scope of this article; it is explored in greater detail in a separate article that advances an innovative sentencing framework for punishing atrocities.[83] Nevertheless, it may be helpful to provide some brief reflections here on how existing ICL sentencing provisions and factors can be interpreted to account for the enabler factor. The relevant statutory language is found in the basic sentencing provisions of all international criminal tribunals, including the International Criminal Court. International judges are called upon to determine a sentence by considering "the gravity of the offence and the individual circumstances of the convicted person."[84] The latter consideration has thus far been underutilised in ICL judicial sentencing analysis. I propose that

[81] CDF Trial Sentencing Judgment (n 15) [84].

[82] *Prosecutor v Plavšić*, Sentencing Judgment, IT-00-39&40/1-S (27 February 2003); See further, Shahram Dana, "The Limits of Judicial Idealism: Should the International Criminal Court Engage with Consequentialist Aspirations?" (2014) 3 Penn State J Int'l Affairs 30.

[83] *See* Shahram Dana, "Reimagining Punishment for Atrocity Crimes: An Innovative Sentencing Framework" (2020) 5(1) Cambridge Law Review (forthcoming).

[84] SCSL Article 19(2); ICC Article 77; ICTR article 23; ICTY Article 24.

the statutory criterion of "individual circumstances of the convicted person" take into account "the role of the accused" as informed by the enabler theory.

ICL judges routinely discuss "the role of the accused" when weighing the appropriate punishment. The sentencing jurisprudence, however, offers a confused and varied treatment of "the role of the accused" as a sentencing factor.[85] Some judges reference "the role of the accused" in their general discussion of "gravity of the offense" and treat it as one of many enumerated "gravity" factors, but the decisions are not consistent in how they conceptualise, integrate, and account for this factor.[86] Other judgments treat "the role of the accused" as part of the accused's mode of liability and participation in the crime.[87] Still, other judgments treat the "role of the accused" as an aggravating factor. Thus, ICL jurisprudence offers three different conceptualisations of "the role of the accused": (1) as a gravity factor; (2) as a factor to assess the nature of individual's participation in a specific underlying crime; and (3) as an aggravating factor. Each of these approaches is problematic.[88] Despite these variations, one thing is clear: "the role of the accused" is an important sentencing factor. But questions relevant to determining its content and influence remain unresolved, making it an important, but unpredictable factor at sentencing.

These problems can be transcended by understanding the "role of the accused" through the prism of the enabler theory. If ICL's conceptualisation of "the role of the accused" is informed by the enabler factor, then "the role of the accused" solidifies as a predictable factor, distinct from the concepts of gravity, modes of liability, and aggravating factors. Moreover, this approach allows the enabler factor to infuse the concept of "the role of the accused" with substance significant and distinctive to mass atrocities, thereby capturing its salience to

85 *ibid*.

86 *C.f.*, *Taylor* Appeals Judgment (n 30); *Prosecutor v Blaškić*, Case No. IT-95-14-A, Appeal Judgment (29 July 2004) [683]; RUF Trial Judgment (n 14) [40]; CDF Trial Sentencing Judgment (n 15) [33]; *Prosecutor v. Aleksovski*, Case No. IT-95-14/1-A, Appeal Judgment (24 March 2000) [182]; *Prosecutor v Brma*, Case No. SCSL-04-16-T, Trial Sentencing Judgment (July 19, 2007) [19]; *Prosecutor v Furundzija*, Case No. IT-95-17/1-A, Appeal Judgment (21 July 2000) [249]; *Prosecutor v Delalic* ("Ćelebići *Case*"), Appeal Judgment (20 February 2001) [731].

87 This ambivalent treatment of "role of the accused" traces its origins to the early jurisprudence of the ICTY. See, *Furundzija* (n 144) [249]; *Blaškić* (n 144) [683]; *Aleksovski* (n 144) [182]; Ćelebići *Case* (n 144) [731].

88 *See* Shahram Dana, "Reimagining Punishment for Atrocity Crimes: An Innovative Sentencing Framework" (2020) 5(1) Cambridge Law Review (forthcoming).

situations of systemic criminality. This approach requires a shift in the discourse currently found in sentencing judgments.[89]

V. Conclusion

In punishing atrocities, judicial sentencing narratives are too preoccupied with justifying their outcomes almost exclusively in terms of "gravity", a rhetoric that has arguably been overplayed and is at risk of having a watered-down effect. The dominance of "gravity of the offence" is crowding out deliberation of meaningful factors crucial to atrocity trials and is stifling the development of the law of atrocity punishment. This article argues for a departure from this approach to punishing atrocities, which has largely transposed domestic sentencing practices applied to ordinary crimes. It offers a novel approach to punishing atrocity crimes: *the enabler theory*. The role of the accused in enabling the milieu or situation in which atrocity crimes spawn has thus far been under-considered at sentencing. I argue that because atrocity crimes are typically the outcome of large-scale violence or armed conflict, ICL punishment must account for a convicted person's role as an enabler.

Even though the enabler factor appears to have a silent influence on sentencing at the SCSL, judges do not explicitly account for the enabler factor in their sentencing analysis. They should. The enabler factor will bring greater coherence and clarity to their reasoning and sentencing allocations. For example, judges justified Charlies Taylor's hefty 50 year prison sentence by simply asserting that his case was "unique." But what precisely is "unique" about Charles Taylor's criminality? The trial chamber never adequately explains this and consequently, several observers criticise Taylor's punishment as too harsh and unjustified go around the court's own sentencing principles.[90] The enabler theory closes this explanatory gap, elucidating why Taylor's sentence is not excessive for the atrocities he enabled while he was a sitting Head of State.[91] Rather than unsatisfyingly ending their sentencing analysis dependent on the "unique circumstances of a case", the judges could have returned to the findings in their judgment where they determined that Taylor had enabled the RUF/AFRC in the conflict and the

[89] The innovation required in interpreting existing ICL statutory criteria and sentencing factors is argued in detail *in* Shahram Dana, "Reimagining Punishment for Atrocity Crimes: An Innovative Sentencing Framework" (2020) 5(1) Cambridge Law Review (forthcoming).

[90] See *e.g.* Heller, *Taylor Sentence* (2013) (n 29) 835–40; Drumbl, Punishing Heads of State (2012) (n 26).

[91] See above Section III.A.

atrocities they committed.[92] A Head of State's wrongdoing that enables large scale violence and atrocities in another country is wrongdoing that must be captured in his punishment. The enabler theory here provides the judges with a justification well rooted in the purpose of atrocity trials and international law.

Likewise, the enabler theory explains differences in sentencing outcomes that gravity alone cannot account for. For example, the enabler factor explains why Sesay, the top RUF commander and an enabler of the conflict and atrocities, received double, or in some cases even triple, the punishment that his subordinates received for the same crimes, even though the subordinates directly committed the killings and Sesay's responsibility was based on an omission—his failure to prevent the subordinate from committing the crime or to punish him afterwards. For his omissions, Sesay received twice or three times the punishment that the direct perpetrators of the crime received. The enabler factor explains why his sentence is appropriate, notwithstanding the fact that his criminal liability is based on an omission. Furthermore, the enabler theory is also helpful in expressing why a group of defendants are receiving lower sentences even though they have committed serious crimes. The CDF fighters were sentenced to substantially less years of imprisonment that the RUF and AFCD soldiers (and Charles Taylor). The average trial sentence for the former group was seven (7) years of imprisonment; whereas the average sentence for the RUF and AFCD perpetrators was forty-four (44) years. The trial judges justified these very lenient sentences on the grounds that the CDF fighters were the "good guys" fighting a just war.[93] The justification, however, runs counter to the ethos of international humanitarian law and international criminal justice. The enabler theory, on the other hand, offers an explanation for the lower sentences that is more congruent with the principles and values of international law.

The enabler theory offers several advantages when punishing atrocity crimes. By departing from a sentencing discourse that narrates atrocity punishment entirely in terms of gravity, my approach allows a space for important nuances to be communicated and expressed in the ICL sentencing process. It also closes the gap between judicial narratives about atrocities and sentencing outcomes. Presently, the normative expressions are compromised under an exclusive reliance on hyper "gravity" imagery that far outpaces and overshadows the actual quantum of punishment. Judicial narratives tirelessly confront the reader about the gravity

92 *Taylor* Trial Judgment (n 5) [5834], [5835], [5842], [6913]–[6915] (finding that Taylor "enabled the RUF/AFRC's Operational Strategy" and "supported, sustained and enhanced the RUF/AFRC's capacity to implement its Operational Strategy.")

93 See above Section III.C.

of the offence and the monstrosity of the perpetrator's crimes.[94] Yet, the final sentences are underwhelming in the face of such explosive rhetoric. This method of expression surrenders too much. Gravity needs to yield its monopoly as an explanatory tool for punishing atrocities. The enabler factor allows the judges to speak the language of gravity to acknowledge the harms suffered by the victims and yet explain why every atrocity perpetrator is not being sentenced to life imprisonment or a lengthy prison term.

Thus, it brings clarity to the accused's moral culpability and makes punishment more communicative. For example, it can account for why individuals whose crimes are of comparable gravity, like Sesay, Kellon, and Gbao, received substantially different sentences. Victims can better understand why the perpetrator of crimes against them received a lesser penalty than another perpetrator, without ICL messaging to them that their suffering is not important or that the crimes against them were not of significant gravity. In this way, the enabler factor optimises the expressive capacity of ICL punishment by offering a clear pathway linking justice to the offender's criminal culpability, to the purposes of international trials, and finally to the sentence itself. Additionally, the enabler factor allows ICL judges to account for a convicted person's contribution to enabling a situation erupting in atrocity crimes. Thus far, moral culpability for this wrongdoing in mass atrocities is not adequately accounted for in the sentence, if at all.

When ICL sentences are understood through the prism of the enabler factor, greater congruency is achieved between judicial narratives about atrocity crimes and the quantum of punishment. This does not diminish the role of gravity but rather brings clarity and transparency to punishing atrocities. Unfortunately, these benefits and the opportunities to optimise the expressive capacity of atrocity trials remain unseized because gravity presently monopolises the narratives in sentencing judgements while the enabler factor, although influential, remains hidden. It is time for the enabler factor to surface.

94 DeGuzman, "Harsh Justice for International Crimes" (n 101) 17–24.

Cambridge Law Review (2019) Vol IV, Issue ii, 105–137

In the Light of Different National Circumstances: Equity under the Paris Agreement

Elkanah Oluwapelumi Babatunde*

I. Introduction

The Paris Agreement is the latest addition to the body of treaties governing climate change. Among other things, the Agreement seeks to revolutionise the principle of equity and common but differentiated responsibilities (CBDR) by providing that it shall apply "in the light of different national circumstances." The Paris Agreement prides itself as a progressive development in the distribution of the costs and benefits of climate change.

Although equity and CBDR have earlier been adopted as guiding principles under the United Nations Framework Convention on Climate Change (UNFCCC), the Paris Agreement adds that the CBDR principle shall apply "in the light of different national circumstances." This addition was a result of agitations around the continued relevance of differentiated responsibility given the emergence of some erstwhile developing countries as economic powers and significant emitters of greenhouse gases. The addition was adopted as a way of narrowing the broad dichotomy of countries into developed and developing countries, where the former took up substantive responsibilities, and the latter were beneficiaries of climate funds. It was to ensure that developing countries with enough capacity to take action would not be excluded from such responsibilities based on the application of equity and CBDR.

This paper argues that this addition was unnecessary and merely cosmetic. It shows an inchoate understanding of the principle of equity and CBDR has expressed in the UNFCCC. The notion of taking national circumstances into account in the distribution of climate change responsibilities is implied in equity and CBDR. Instead, I argue that the intention of the Paris Agreement to flesh out respective capabilities in the distribution of commitments is achieved not

* Elkanah O Babatunde is a Ph.D. Candidate at the University of Cape Town, South Africa.

through the addition of this phrase but through other provisions in the Agreement. The caveat 'in the light of different national circumstances' is mere tautology and contributes nothing to both the jurisprudence and the practice of equity in international law.

II. Domestic law origins of equity

Equity is perhaps one of the oldest legal concepts, and its use and manner of application have differed from one legal system to another. The concept has also been utilised by politicians, moralists, political philosophers, among others. Equity has thus gained a wide range of meaning, and its precise meaning would vary depending on the context within which it is used. Moralists and philosophers would differ from politicians and diplomats in the use of the word. Its use in different treaties, by the international court of justice or tribunals, within the General Assembly or the Security Council would also differ from what is meant when it is used in domestic legislation or within the strict common law sense.[1]

The history of equity, which shall be explored below, leads to two different understanding of the concept. One is the idea that equity is the prerogative of judges used as a means of correcting or mitigating the harshness that will result in the direct application of the law to a specific situation. The other school views equity as a rational decision-making process which forms part of natural law and therefore sees equity as a source of law rather than distinct from the law.[2] Equity is thus seen as a part of the judicial adjudication process.[3] This disparity has dictated several legal and political debates on the concept of equity such as the positivist-naturalist debates, the form and role of equity in international and domestic law, the North-South debate, and the extent to which equity should be adopted in law courts and other legislative or judicial process. The following sections would highlight some of these differences and how they have impacted the use of the concept in international law on climate change.

Equity, like many other legal concepts, does not have a linear history or a crystal-clear definition. However, much of what is known as equity today emerged from Aristotle's famous treaties, *Nicomachean Ethics*.[4] According to him, equity and justice were not genetically different. He explains this by pointing out that if equity were different from justice, either of them would be good and the other evil, and it

[1] Christopher Rossi, *Equity and International Law: A Legal Realist Approach to International Decisionmaking* (Transnational Publishers 1993) 3.

[2] *ibid* 21.

[3] Wolfgang Friedmann, 'The Uses of General Principles in the Development of International Law' (1963) 57 Am J Int'l L 279, 287.

[4] Aristotle, *Nicomachean Ethics* (Roger Crisp edn, Cambridge University Press 2000) 99–100.

would be odd that equity receives our commendation if it is evil. If they are both good, it means they are not genetically different. Aristotle, however, maintains that although they are both good, what is equitable is superior when compared with 'strict justice', strict justice being the strict adherence to the letter of the common law.[5] The need for equity arises due to the nature of common law rule, which is a general statement of what is right or just. Although legal adjudication will usually be on general terms, there would be situations where the application of 'an otherwise universally valid rule of law' would result in injustice.[6] In such cases, the justice of the situation falls outside the general rule.[7]

Equity will thus create an exception to the general rule by taking a more definite and specific judgement of what is just in that particular situation. In equity, the law is adapted to the facts of individual cases but must conform to the general notions of law and justice in doing so.[8] Equity was thus not contrary to, but supplementary to common law justice.[9] It is an efficient way of ensuring that justice which is the objective of the common law, is achieved. Equity, as espoused by Aristotle, is thus an exception created to rectify the ills of common law in specific cases and remained an integrated part of the legal system.

A. Equity under the roman law

The concept of equity gained primacy in Roman law through the adjudication of the praetors (magistrates). It had developed as a set of small and inconsistently applied, procedural remedies applied by the Roman magistrates known as praetors to matters of property rights and contracts.[10] It gave room for equity to serve as a tool for remedying situations among individuals who were otherwise politically equal. The magistrates modified the *ius civile* by introducing equitable edicts. These edicts gave remedies to litigants such as a widow whose husband had died intestate and would thus have had no cause of action under the civil law. The magistrates thus developed the jus honorarium to make up for cases where it was necessary to achieve justice even though the parties may not have followed the legal formalities that should have given rise to such rights.[11]

[5] *ibid*.

[6] Anton-Hermann Chroust, "Aristotle's Conception of Equity" (1942) 18(2) Notre Dame Law Review 123.

[7] Rossi (n 1) 22.

[8] Michael Akehurst, 'Equity and General Principles of Law' (1976) 25(4) Int'l and Comp Law Quarterly 801; Chroust (n 6) 124.

[9] Aristotle (n 4).

[10] Justice Margaret White, "Equity: A General Principle of Law Recognised by Civilised Nations?" (2004) 4(1) QUTLJJ 105; Rossi (n 1) 28.

[11] *ibid* 105.

However, equity under Roman law never developed beyond a set of imprecise remedies borrowed from natural law.[12]

B. Equity in the civil law system

The development of equity within the civil law system followed a significantly different route when compared to its development under the common law. While the common law system had allowed for judges-made law under the courts of law and equity, civil law had no such equivalence. On the other hand, civil law followed a stricter code-based system of adjudication where certain officials codified what was believed to be the law. These legal scholars known as the Glossators and Commentators were responsible for bringing together the law in written form and ensured consistency.[13] Equitable principles were included in these collections of the law in order to make up for the strictness of the code system. This codification helped to give the principles full effect in the adjudication of the law without having the dual nature which it had under the common law system.[14]

Various civil codes have since then developed, which make provision for moral justice as part of the legal adjudication process further fusing the direct application of legal process and equitable principles.[15] Provisions, therefore, abound which provides that where a statute requires a judge to decide according to his estimation of the circumstances, or to proper reasons, he must make such a decision in accordance with 'justice and equity.' Others make use of the phrase 'doctrines of natural law'[16] or 'natural equity'[17] or 'doctrines of equity'.[18]

C. Equity under english common law

In England, equity gained prominence through its application in individual cases. Even though judges never intended to establish a distinct set of rules, equity eventually developed into "an independent source of law… a new system of law."[19]

[12] Rossi (n 1) 31.

[13] *ibid* 38.

[14] Ralph Newman, "Equity in Comparative Law" (1968) 17 (4) Int'l and Comp Law Quarterly 830–1.

[15] Rossi (n 1) 38–9.

[16] The Austrian Civil Code, section 7; The Argentine Civil Code, article 7; The French Civil Code, article 21; The Constituent Assembly Law of Ghana, article 4.

[17] The Civil Code of Colombia, article 32; The Ecuador Civil Code, article 17–18; The Honduras Civil Code, article 20.

[18] The Puerto Rico Civil Code, article 7.

[19] Gustav Radbruch, "Justice and Equity in International Relations" in Norman Bentwich, *et al* (eds), *Justice and Equity in the International Sphere* (Constable & Co 1936) 2.

Equity had developed as part of the development of the common law system.[20] The medieval historian, GB Adams, asserts that there was no clear distinction between equity and common Law until the fourteenth century. Both systems operated together in terms of the institutions that administered these rules and the functions of the rules.[21] Equity developed as an attempt by the Lord Chancellor to ensure security and justice in the kingdom when common law had failed or proved inadequate in the circumstances.[22] The procedure for equity usually began as 'petitions for grace' asking the king to interfere with securing justice where the regular law failed to do so.[23] Lord Woolsey commenting on the need to apply equity in some instances made it clear that 'the king ought to… mitigate the rigour of the law, where conscience hath the most force; therefore, in his royal place of equal justice, he hath constitute a chancellor, an officer to execute justice with clemency, where the rigour of law opposes conscience.'[24]

The development of equity was further aided by the principle of *stare decisis*, which ensures that judicial decisions are consistent and predictable.[25] The application of stare decisis helped not only to concretise equity as part of the legal system but to cure one of the significant criticism of equity which was its susceptibility to the whims of whoever was chancellor. The separation in the administration of equity has led to criticisms that equity is outside the administration of law and reliance on it was a reliance on concepts that are external to the law. This confusion persists to this day and has had an impact on the application of equity in international law.[26]

Equity under the common law thus became a means of reconciling legal certainty (which is necessary for social order) with justice which is essential in specific situations: the attempt to ensure that justice is achieved amidst the quest for legal certainty.[27] Equity corrects what would have otherwise been failures of justice due to the incapacity of strict law to adjust to specific situations.[28] Unlike common law which concerns itself strictly with rights and wrongs according to the law, equity takes a more comprehensive view of right and wrong in its approach to justice. It

20 George Adams, "The Origin of English Equity" (1916) XVI (2) Columbia Law Review 88.

21 *ibid* 89.

22 Richard Hedlund, "The Theological Foundations of Equity's Conscience" (2015) 4 Oxford Journal of Law and Religion 123.

23 Adams (n 20) 91.

24 Shyamal K Chattopadhyay, "Equity in International Law: Its Growth and Development" 5 Ga J Int'l & Comp L 381.

25 Rossi (n 1) 36.

26 *ibid* 32–3.

27 Ralph Newman, "An Analysis of the Moral Content of the Principles of Equity" (1967) 19(1) Hastings LJ 150.

28 *ibid* 151.

considers the unique situation of the parties involved and what relationships may exist between them to make the case worthy of special consideration.

It is thus clear that whether within the old Greek tradition, Roman, common or civil law system, equity was recognised as a concept for the mitigation of general legal principles in order to achieve justice in specific cases. Judges and litigants relied on these concepts frequently to plead their case where it appeared that the application of strict legal principles would not avail them. The aim was not to derogate or undermine established legal principles and the rights which they may have created but to ensure that these principles do not serve as obstructions to justice in some instances. It was a mechanism for marrying the quest for legal certainty with the need for justice.

III. Equity in international law

Although the origins of the equity principle can be traced back to various domestic legal systems, equity has also become an enshrined concept within the international legal system.[29] Hugo Grotius was one of the thinkers who first emphasised the application of equity in international law. He stated that "… upon this principle [equity] all wills and treaties ought to be interpreted. For as all cases could neither be foreseen nor expressed by the lawgiver." He concludes that it is, therefore, necessary to vest in the judiciary the power to make exceptions in the interpretation of the law in some instances.[30]

Equity was first introduced into the international legal system through the practice of international arbitral tribunals.[31] From early treaties such as the Jay Treaty of 1794 to the popular *Alabama Claims Arbitration* of 1871–1872, several treaties followed between the late 19th century and the first half of the 20th century which provided for arbitration as a means of solving disputes between states.[32] Principles of natural law[33] were introduced into the international legal system through these arbitral decisions. The arbitrators employed natural law principles in stemming down the strictness of the treaties and legal agreements which they had to apply in their arbitration. The 1794 Jay Treaty provided that the Commissioners "decide the claims in question according to the merits of the several cases, and to justice, equity and the law of nations." The 1892 *Hero, Nutrias*

[29] Friedmann (n 3) 287; Ruth Lapidoth, "Equity in International Law" (1987) 81 Proceedings of the Annual Meeting *(American Society of International Law)* 139.

[30] Hugo Grotius, *The Rights of War and Peace* (Campbell translation 1901) 26.

[31] Clarence Jenks, *The Prospects of International Adjudication* (Stevens 1964) 319; Rossi (n 1) 87.

[32] Rossi (n 1) 42.

[33] Natural law is that claim which states that there is a set of moral standards to which the law must conform if it is to be regarded as just.

and San Fernando steamships cases between the United States and Venezuela were also to be decided "in accordance with justice and equity and the principles of international law".

This practice continued into the 20th century with the *Orinoco Steamship Arbitration 1910*, where the tribunal was to adjudicate "in accordance with justice and equity". It was also provided by a treaty that "the principles of law and equity" be applied in deciding the *Norwegian Shipowners' Claims (Norway v United States)*.[34] The tribunal held that this implies that the tribunal would be within its jurisdiction to apply equity where positive law is lacking in an area of the dispute between the parties.[35] In this case, equity was understood by the tribunal as a general principle of justice.[36]

The tribunal in the *Cayuga Indians case* noted that where the application of positive law would lead to an abnormal result, recourse must be made to "generally recognised principles of justice and fair dealings."[37] The tribunal explained that the purpose was to avoid reaching inequitable results.[38] The Gran Chaco conflict concerning the border between Bolivia and Paraguay (1932–38) was also settled by an arbitral award based on equity and good conscience. This award presents a compelling case as the decision to rely on equity had been based on the undesirability of relying on written law because it would have led to an undesirable result.[39]

In the *Norwegian Ship Owners Claim Case*, the United States government had requisitioned some ships during World War 1, including some being built for Norwegian citizens. The Norwegian government, therefore, requested for payment on behalf of its citizens. The American government however, refused, and a dispute therefore arose. In applying 'law and equity' to the case, the arbitrators explained that "… these words are to be understood to mean general principles of justice as distinguished from any particular system of jurisprudence or the municipal law of any state."[40] Also, the arbitral tribunal in the *Indo-Pakistan Western Boundary (Rann of Kutch)* case had no express consent of the parties to apply equity. The arbitrators, however, concluded in the preliminary discussions that "as both Parties

[34] *Norwegian Shipowners Claims* (*United States of America v Norway*) Reports of International Arbitral Awards 13 Oct 1922 vol 1 307, 310.

[35] *ibid* 331.

[36] *ibid*.

[37] *Cayuga Indians (Great Britain v United States)* Reports of International Arbitral Awards 22 January 1926 vol VI 174, 180.

[38] *ibid*.

[39] *Dispute between Bolivia and Paraguay* (1934) Report of the Chaco Commission 45.

[40] *Norwegian Shipowners Claims* (n 34).

have pointed out, equity forms part of international law; therefore, the parties are free to present and develop their cases with reliance on principles of equity."[41]

In the *Meuse Case* before the PCIJ, Judge Manley Hudson maintained this position and stated that: "… principles of equity have long been considered to constitute a part of international law, and as such they have often been applied by international tribunals… Article 38 of the Statute expressly directs the application of 'general principles of law recognised by civilised nations', and in more than one nation, principles of equity have an established place in the legal system. The Court's recognition of equity as a part of international law is in no way restricted by the extraordinary power conferred upon it 'to decide a case *ex aequo et bono*, if the parties agree thereto'... It must be concluded, therefore, that under Article 38 of the Statute, if not independently of that Article, the Court has some freedom to consider principles of equity as part of the international law which it must apply."[42]

Judge Anzilotti reinforced this position stating that the principle of equity is "so universally recognised, that it must be applied in international relations also."[43] The ICJ noted in *Military and Paramilitary Activities in and against Nicaragua ([Nicaragua v United States of America]* that anyone is yet to object to this position.[44] In the *Frontier Dispute Case*, a dispute had arisen as to the exact delimitation of the boundaries of the Burkina Faso and Mali. Mali had urged the court not to apply Article 38(2). Instead, the court should apply a "form of equity which is inseparable from the application of international law".[45] It is worthy to note that Burkina Faso never objected to this conception of equity as inseparable from the application of international law. In its decision, the ICJ acknowledged that equity was inextricably tied to the idea of justice and held that it would apply that form of equity which ensures a proper interpretation of the law.[46] The court here was taking a typical Aristotelian position, employing equity to ensure that the interpretation of the law does not lead to absurdity. Thus, showcasing that equity had become a part of the

41 *The Indo-Pakistan Western Boundary (Rann of Kutch) between India and Pakistan* Vol. XVII Report of Arbitral Tribunals 19 February 1968 1, 11.

42 *Diversion of Water from the Meuse (Netherlands v Belgium)* Judgment No. 25 28 June 1937 PCIJ Ser A/B, No 70, [322].

43 *ibid* 211.

44 *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America)* (1986) ICJ Rep para 269.

45 *Case Concerning the Frontier Dispute (Burkina Faso v Mali)* ICJ Reports 1986, 22 December 1986, [27].

46 *ibid* 149.

international legal system even without express consent as contemplated in Article 38 (2) of the ICJ Statute.

The principle of equity was also central to the ICJ decision in the *North Sea Continental Shelf* cases. The case centred on the delimitation of the continental shelf between the Federal Republic of Germany and Denmark on the one hand, and between the Federal Republic of Germany and the Netherlands on the other hand. The court was saddled with the responsibility to determine the legal principles upon which the delimitation was to be carried out. Rejecting the application of the equidistance principle as provided for by the Geneva Convention, the court held that the shelf was to be delimited by agreement between the parties and in accordance with equitable principles.[47] The court, however, cautioned that the surrounding circumstances be taken into account in the application of these principles. Equity was not to be applied as a matter of abstract justice.[48]

In reaching this decision, the court was careful to distinguish its application of equitable principles (*intra legem*) from *ex aequo et bono*.[49] The court made it clear that the introduction of equity, in this case, was not tantamount to acting outside the purview of the law. The objective justification for the application of equity was within the law. Express consent of the parties as provided for in Article 38(2) of the Statute of the ICJ was therefore unnecessary in the present case. The court emphasised that its duty to dispense justice means that its decision must "find its objective justification in considerations lying not outside but within the rules." The application of equity must be founded within the rule of law, thus ensuring that judges do not exercise an unrestrained discretion in determining cases. The court held that equity was a general principle of law applicable to international legal decision making.[50]

This decision has, however, been criticised. Friedmann, for example, noted that "by rejecting the criteria laid down in the convention and other documents, the court, in effect, was giving a decision *ex aequo et bono*, under the guise of interpretation".[51] Another scholar, Janis criticised the ICJ's opinion (per Judge Hudson) as too Anglo-American and warns against adopting an exclusively Anglo-American interpretation of equity in international law. He points to current

[47] *North Sea Continental Shelf Cases, (Germany v. Denmark; Germany v. Netherlands*) 1969 ICJ 53.

[48] *ibid* 85.

[49] While the application of equity *intra legem* refers to the application of equity founded upon the law, *ex aequo et bono* refers to that application of equity not prescribed by the law and hence only applicable by consent of parties to the case.

[50] *ibid* 88.

[51] Wolfgang Friedmann, 'The North Sea Continental Shelf Cases: A Critique' (1970) 64 Am. J. Int'l L. 236. It is difficult to find a case where parties have consented to the ICJ's exercise of ex aequo et bono under Article 38(2) of the Statute of the ICJ.

international practices and especially the views of third world scholars which see equity as distributive, rather than merely discretionary, justice.[52] It must, however, be noted that while these scholars may have disagreed on the specific interpretation of equity, they do not dispute that there is a sense in which the application of equity is required in international legal decision making. Equity is generally accepted as an established principle of international law.[53]

IV. Contesting claims of equity

It is clear from the analysis above that the principle of equity has become established as a part of the international law jurisprudence. It is, however, dangerous to assume that language in international law has a single or permanent meaning.[54] The sense in which a concept or principle is used in international law differs across various scholars, international lawyers and practitioners. These concepts also evolve continuously from time to time and may have several meanings across various contexts. This challenge is particularly apparent with the use of equity. Its application (or non-application) within international law is made more difficult by economic and social differences between countries and across regions of the world.[55]

Concepts and principles within international law have been fraught with varying interpretations due to the different legal, political, doctrinal and even economic traditions that converge in the international legal system. Thus, even seemingly simple terminologies have proved difficult to give precise meanings and interpretations, and this has been even more complex with the principle of equity.[56] Having established that equity had become a common principle adopted within the international legal system, it is necessary to examine the sense in which the concept has been used within the system.

It is necessary to state from the outset that although international case law is replete with the application of equity, judges on these international tribunals generally do not state that they are applying equity or provide a direct equity analysis.[57] Lapidoth explains that the practice has been to refer to a particular principle without referring to its equity origins.[58] Thus, it is commonplace to see such principles as proportionality applied in rulings on the use of force, estoppel,

[52] Mark W Janis, 'The Ambiguity of Equity in International Law' (1983) 9 Brook. J. Int'l L. 26-29.

[53] Friedmann (n 3); Lapidoth (n 29) 139.

[54] Janis (n 52) 7.

[55] Peter Thacher, "Equity under Change" *Proceedings of the Annual Meeting* (American Society of International Law 1987) 134.

[56] Janis (n 52) 7.

[57] Lapidoth (n 29) 140.

[58] *ibid* 144.

etc. without a direct reference to equity. This has, therefore made it difficult to ascertain the specific meanings or implications of the application of equity in international law.

A. Equity as distributive justice

By the turn of the twentieth century and with much of the global south gaining independence and joining the global system and international law, agitations began about the need to restructure the existing international social, political and economic structures. Newly independent African, Asian and Latin American countries began to argue for a new economic order where the distribution of global wealth will be more evenly shared among developed and developing countries of the world.[59] The use of equity in the call for a new international economic order found its root in the economics and politics of decolonisation.[60] Scholars and practitioners alike began to clamour for a more just distribution of global wealth, and they found equity as a useful normative tool for pushing for this position within the international legal system.

Equity, even from its common law origins, had the characteristics of emphasising the need to do justice by taking cognisance of the peculiar facts of each case. The distinguishing characteristic if equity in the judicial decision-making process is the freedom which it gives the decision-maker to make decisions without the restrictions of established legal rules and principles. This characteristic makes equity suitable for contexts where there are competing interests, such as access to shared resources, but very little law designed to address such interests or inequality.[61] It was, therefore, a flexible tool for advocating for justice in the global distribution of wealth and for ensuring that members of the global South have access to the basic necessities. Third world scholars and practitioners thus began to transform the uses of equity from the discretionary sense in which it had been widely known into a more substantive concept for ensuring fair distribution of the world's resources.

This agitation led to the adoption of the UN General Assembly resolution on the Establishment of the New International Economic Order. This resolution provides that states work together for an order "based on equity, sovereign equality, independence, common interest and co-operation among all States, irrespective of their economic and social systems which shall correct inequalities and redress existing injustices."[62] The objective of this order was to eliminate the gap between

[59] Janis (n 52) 17.
[60] Janis (n 52) 1.
[61] Vaughan Lowe, "The Role of Equity in International Law" (1988–1989) 12 Aust YBIL 73.
[62] GA Res. 3201 (S-VI) UN GAOR Supp (No. 1) 5 UN Doc A/9556 in 13 ILM (1974) 715.

developed and developing countries and ensure economic development in the latter. It sought to reduce global inequality and ensure prosperity for all.[63] The UN Conference on Restrictive Business seems to have adopted the same position. The conference declared that "in order to ensure the equitable application of the Set of Principles and Rules on Restrictive Business, states should take into account in their control of restrictive business practices the development, financial and trade needs of developing countries…" The resolution obliged developed countries to promote local industries and economic development within these developing countries.[64] Equity in this essence was a form of distributive justice, ensuring that global wealth was distributed to regions of the world that previously had little or nothing of this wealth.

Equity was thus taking a new definition under the new international economic order. Unlike previous use of equity under the arbitral tribunals and the ICJ, equity here was not discretionary. It was an obligation, founded upon the principle of justice and fairness, to reduce the developmental gap between the nations of the world. It implied the sharing of the benefits of technological progress, a just model for the determination of prices of raw materials and primary commodities exported by developing countries and those imported by them, and generally ensuring a just and equitable international division of labour.[65] A similar reference was made to equity in the Programme of Action on the Establishment of a New International Economic Order and in the Charter of Economic Rights and Duties of States.

However, these references to equity do not impose any legal character on the concept. They were more or less political statements of aspiration for a just distribution of global wealth between developed and developing countries without any legal implication. This is especially so when one juxtaposes these notions of equity with the traditional Aristotelian sense in which equity is used in the Common law systems: a form of discretional powers to be used by a judge or arbitrator in specific cases. Janis responds to this kind of query by noting two things. First, subsequent practice as shown that reference to equity does indeed imply some measure of a legal rule. Secondly, such claims fail to take into cognisance the

[63] Janis (n 52) 17.

[64] UN Doc TD/RBP/CONF./10 (May 2, 1980) in 19 ILM 813 at 814.

[65] GA Res 3201 (S-VI) UN GAOR Supp UN Doc A/9556 in 13 ILM (1974) 715, 716; GA Res 3202 (S-VI) UN GAOR Supp (No 1) 5 UN Doc A/9631 (197) 14 ILM 251, 252.

existence of other legal principles whose application give legal quality to these references to equity.[66]

Bedjaoui's explanation of equity in his treatise *Towards a New International Order*, follows the distributive justice theory.[67] Bedjaoui, a former judge of the ICJ, starts his analysis by arguing that Europe and the United States are largely responsible for the inequality in the global distribution of wealth. He contends that this group of countries have exploited resources at the expense of the Third World, and they are only able to live in wealth at the expense of the developing world. He further argues that the present international legal system, under the guise of neutrality, has helped to consolidate the gains of this unequal economic relations if not to create it.

Despite his incrimination of the international legal system, he argues that the system can also be a tool for correcting this global inequality. This, however, would imply that international law is freed from its "paralysing formalism to steer it towards a nobler, more humane and more essential goal- the promise of development."[68] Relying on the general idea of equity as that process through which a judge reaches a decision which ensures justice given the peculiar circumstances, Bedjaoui sees equity as that avenue through which the inequalities in the distribution of global wealth may be corrected. He, therefore, concludes that international law "must give great prominence to the principle of equity. In doing so, it must keep the objective in view which consists of reducing and, if possible, eradicating the gap that exists between a minority of rich nations and a majority of poor nations."[69] He thus sees equity as a norm to be deployed for purposes of redistributing global wealth and correcting existing inequalities in the level of development between various countries.

Although Janis agrees with Bedjaoui on the necessity for equity, he questions the need to transform the international legal system radically. Janis argues that the existing international legal system had inherent flexibility which could be adopted to achieve the equitable ends for which Bedjaoui argued. There is, therefore, no need to jettison or cause a radical change in the traditional legal system as Bedjaoui seem to be suggesting.[70] Janis also disputed Bedjaoui's claim that international law was too divorced from principles of natural law. In support of this contention, Janis gives an index of scholarly writings dealing with the relationship between

[66] Janis (n 52) 19.

[67] Mohammed Bedjaoui, *Towards a New International Economic Order* (Holmes & Meier 1979).

[68] *ibid* 63.

[69] *ibid* 119.

[70] Mark Janis, "Towards a New International Order by Mohammed Bedjaoui" (1983) 6 B C Int'l & Comp L Rev. 357.

international law and equity.[71] While Janis was right in pointing out that there were indeed certain elements of natural law in international law and Bedjaoui was not entirely accurate with his claim, the texts he provides in support of his argument miss the main points of Bedjaoui's conception of equity.

These texts merely discuss equity as a form of discretionary power to be exercised by the international arbiter or judge in reaching decisions in a dispute. Bedjaoui was, however, referring to equity in a different form, not merely as a procedural or discretionary power of a judge, but as a tool for distributive justice. Haq takes a similar position on equity when he argues for an obligatory transfer of resources from developed to developing countries in order to achieve "economic equity in the global context."[72] Janis adequately summarises these notions of equity as "a form of distributive justice, aimed to meet the needs of developing countries."[73]

B. Equity as discretionary justice

Ian Brownlie and Lauterpacht, two prominent international law scholars, however, disagree with the distributive justice schools of thought on the role and interpretation of the concept of equity within the international legal system. Contrary to the notion of equity as a corrective measure to address the global economic inequality between countries, these scholars view the role of equity in the international legal system as a tool to ensure a more flexible international system, especially where specific cases cannot be fully adjudicated by strict rules.[74] They both insist on equity as the ability of the judge to exercise some form of discretionary power in the pursuit of justice.

Brownlie rejects the idea of equity as a tool for solving global problems such as global economic inequality. According to him, equity offers "little but disappointment as a tool for solving these sophisticated problems."[75] Equitable principles in this context are according to him "no more than a bundle of impressionistic ideas…"[76] Similarly, Lauterpacht restricts the invocation of equity in international law to that discretionary powers to be exercised by judges, and when a judge applies this subjective or discretionary principle, parties must be

[71] *ibid* 359.

[72] Inamul Haq, "From Charity to Obligation: A Third World perspective on Concessional Resource Transfers" (1979) 14 Tex. Int'l LJ 406.

[73] Janis (n 52) 22.

[74] *ibid*.

[75] Ian Brownlie, "Legal Status of Natural Resources in International Law (Some Aspects)" (1979) 162 Recueil des Cours 287.

[76] *ibid*.

given the opportunity to make submissions on such principles.[77] Equity within international law was thus for him a means of ensuring a more flexible and responsive international law.[78]

Another scholar of the Western tradition, Lowe, argued that the application of equity is not necessary for international law since there will always be other techniques which could be applied to fill whatever gaps may exist in international law.[79] He argues that it will always be possible to do justice in international law without recourse to "a distinct normative source in the form of equity."[80] He contends that the equitable application of the law can achieve the end sought by advocates of equity.[81] A tribunal merely needs to interpret existing law in a manner that does justice, rather than introduce extra-legal factors into the decision making process. Koskenniemi holds similar views and sees the invocation of equity as unrealistic attempts to revive natural law.[82]

Lapidoth, however, cautions that the subjectivity of equity is dangerous for international law. She contends, and rightly so that the conception of equity is dependent on an individual's "ethical environment". She continues that the fact that this will be different for each actor in the international legal system implies that the qualities of the law such as 'generality, clarity, certainty and predictability' will be lost.[83] The subjectivity of equity would make a decision which is just to one party unjust to another. Lapidoth, however, concedes that equity is part of the law and must be applied appropriately.[84]

Although an overwhelming majority of Third World Scholars and the Resolution on the New International Economic Order supports the notion of equity as distributive justice, there is no consensus on such interpretation of equity within the international legal system. For example, Janis cautions that equity has been used within other contexts where its precise meaning is not readily deducible.[85] He cites Article 11(7) of the UN 1979 Agreement Governing the Activities of States on the Moon and Other Celestial Bodies which states that the purpose of the regime was to establish "an equitable sharing by all State Parties in the benefits derived from those resources, whereby the interests and needs of the developing countries,

77 Elihu Lauterpacht, "Equity, Evasion, Equivocation and Evolution in International Law" in Proceedings and Committee Reports of the Am Branch of the Int'l L Ass'n (1977–78) 45–6.

78 *ibid* 46.

79 Lowe (n 61) 63.

80 *ibid*.

81 *ibid* 65.

82 Matti Koskenniemi, *From Apology to Utopia: The Structure of International Legal Argument* (CUP 1989).

83 Lapidoth (n 29) 146.

84 *ibid* 147.

85 Janis (n 52) 26.

as well as the efforts of those countries which have contributed either directly or indirectly to the exploration of the moon, shall be given special consideration."[86] Thus, although the agreement uses the word 'equity', the implication in this context is not specific.

Janis suggests that equity within this context can be interpreted both in the discretionary and the distributive justice dimension. If one takes the distributive interpretation, it implies that the benefits from moon exploration would be shared with developing countries based on their needs rather than their direct contribution to the exploration process. On the other hand, a discretionary justice interpretation would imply that parties would negotiate the distribution of benefits at a later time, and while they should take the needs of developing countries into consideration during such negotiations, the guiding rule will be the contribution of each state to the exploration of the moon.[87]

In explaining the uncertainty in the meaning and implication of equity within international law jurisprudence, Janis also examines the UN Convention on the Law of Sea which makes use of the term 'equity' and 'equitable principles' in several contexts with unclear meanings. While the use of equity in the preamble seems clear and refers to the distributive sense of the word (as used in the new international economic order), equity in the delimitation provisions seems to refer to negotiated settlements and flexible decision making (the discretionary sense as argued by Lauterpacht and Brownlie). The other uses of equity within the UNCLOS are mostly uncertain.[88]

Given these uncertainties and the inability of both the discretionary and the distributive justice theory to adequately explain the use of equity within international law, Janis postulates a third theory: equity as measured justice. Equity as measured justice seeks to combine both the discretionary and distributive sense of equity. According to Janis, equity is achieved by taking two steps. The first is to ask whether it is such a case that requires that decision be made on a discretionary basis and secondly, how that power of discretion is to be exercised. In other words, can extra-legal principles be invoked in the decision-making process or the distribution of rights and obligations? If the answer is in the affirmative, what extra-legal principles can be adopted in this process? To this second question, he answers that developmental need or distributive justice becomes one, among other

[86] UN Doc A/34/664 (12 November 1979) cited in MW Janis, "The Ambiguity of Equity in International Law" (1983) 9 Brooklyn J Int'l Law 7, 27.

[87] Janis (n 52) 27.

[88] *ibid* 28–9.

factors which may be adopted. The litmus test is that the decision reached must be proportional to the non-legal standard adopted.[89]

V. Equity and cbdr in international climate change law

Prior to the 1992 Conference on the Environment and Development and the adoption of the UNFCCC, equity had begun to emerge in a wide variety of environmental law treaties. Examples include the 1982 UN Convention on the Law of the Sea which provided for the: realisation of a just and equitable international economic order", for an "equitable solution" in the achievement of the objectives of the Convention and several other references to the principle of equity in one form or the other.[90] Others include the 1979 Agreement Governing the Activities of States on the Moon and Other Celestial Bodies, the Biodiversity Convention with an apparent reference to fair and equitable sharing,[91] the Oil Pollution Convention which provides for the equitable geographic distribution of membership on Executive Committee.[92] These treaties showed an increase in the inclusion of the principle of equity within treaties, especially treaties which concerned environmental issues.

The United Nations Conference on the Environment and Development (Rio Conference) held in 1992 was, however, the first major conference where countries had the opportunity to negotiate and adopt a treaty on climate change. Countries had gathered at this conference backed by the discovery by scientists that human action was leading to a change in the state of the climate and that there was a need to reduce anthropogenic emissions of greenhouse gases. Prior to the conference, the United Nations General Assembly had established the World Commission on Environment and Development (the Brundtland Commission). The Commission had the responsibility to recommend how countries at different stages of development could cooperate to achieve the common objectives of economic development and environmental protection.[93]

In its report, *Our Common Future*, the Commission noted that an ecological crisis would persist as long as poverty and inequity remain endemic.[94] The Commission noted the widening resources gap between developed and developing countries and the dominance of industrialised countries in international rulemaking. According to the Commission, this "inequality is the planet's main environmental problem"[95]

[89] *ibid* 30–3.
[90] Preamble, articles 74, 83 and 161 of the UN Convention on the Law of the Sea.
[91] Articles 1 and 15(7) of the Biodiversity Convention.
[92] Article 22(2).
[93] UN General Assembly resolution A/38/161 19 December 1983 para 8(b).
[94] *Report of the World Commission on Environment and Development: Our Common Future* (1987) ch 2, [4].
[95] *ibid*. 17.

The Brundtland Commission, therefore, concluded that any attempt to address environmental concerns without tackling underdevelopment in some regions of the world would be futile.[96] In other words, global environmental governance would only stand a chance if it is coupled with provisions which ensure economic growth for the developing world.

It was thus clear from its inception that the development of a treaty on climate change would have to deal with issues of equity which had earlier been raised at the 1972 UN Conference on the Human Environment and echoed by the Brundtland Report. The position of equity in climate change law and policy, however, proved contentious due to the nature of climate change. This is due, among other reasons, to the fact that historically, climate change is a consequence of industrialisation and consumption patterns which are most prevalent in developed countries.[97]

Although all countries contribute to the build-up of greenhouse gases, it is generally agreed that developed countries are responsible for the majority of the accumulated greenhouse gas (GHG) emissions.[98] Developed countries are responsible for over seventy per cent of global emissions while developing countries contribute about a quarter even though they have eighty per cent of the world's population.[99] As at the early 2000s, the United States alone was responsible for almost twenty-five per cent of the world's greenhouse gas emissions as compared to 136 developing countries who contribute a total of twenty-four per cent of these emissions.[100] Peter Hayes even went as far as making calculations to deduce how much cost different countries were responsible for as a result of their historic contributions to climate change.[101] Earlier in the 1960s, the Swedish scientist Borgström had explained that many European countries were using the resources of emerging countries to enrich their own countries while leaving these countries

[96] Lorraine Elliott, *The Global Politics of the Environment* (2nd edn, Macmillan 2004) 171.

[97] Karin Mickelson, "Leading Toward a Level Playing Field, Repaying Ecological Debt, or Making Environmental Space: Three Stories About International Environmental Cooperation" (2005) 43 Osgood Hall LJ 150–4.

[98] Marina Cazorla and Michael Toman, "International Equity and Climate Change Policy" (2000) 27 Climate Issue 2.

[99] Jos Olivier, Greet Janssens-Maenhout, Marilena Muntean and Jeroen Peters, *Trends in Global CO2 Emissions: 2016 Report* (NEAA 2006) 34; John Ashton and Xueman Wang, "Equity and Climate in Principle and Practice" in Joseph Aldy, *et al* (eds), *Beyond Kyoto: Advancing the International Effort against Climate Change* (Pew Center on Global Climate Change 2003) 5.

[100] Timmons Roberts and Bradley Parks, "Ecologically Unequal Exchange, Ecological Debt and Climate Justice: The History and Implications of Three Related Ideas for a New Social Movement" (2009) 50(3–4) Int'l J of Comp Soc 393.

[101] Peter Hayes, *The Global Greenhouse Regime: Who Pays?* (Earthscan 1993).

worse off. It is thus apparent that the effects of climate change are disproportionate to the historical responsibility and present capabilities for this phenomenon.[102]

Developing countries argued during the Rio conference that the primary duty to act with respect to environmental protection lies with developed countries.[103] Developed countries owe an ecological debt to developing states, and since they are the major beneficiaries of climate change, they are liable for its consequences.[104] They were concerned that responsibility should be placed on them for a problem which they had played little or no part in creating.[105] Developing countries have consistently emphasised that since the root cause of climate change lies in the environmental effects of industrialisation, much of which has been to the advantage of developed countries.

Industrialised states should, therefore, take the lead in addressing environmental problems.[106] Under the auspices of the G77, developing countries maintained that developed countries had both a "historic and current responsibility" for environmental problems due to their unrestrained exploitation of natural resources in the past.[107] Unlike what was later adopted, the G77 had stated expressly that developed countries bear the "main responsibility for global environmental degradation." The provision of financial and technological resources was, therefore, to be regarded as a form of compensation for this past exploitation of resources which has resulted in global harm.[108]

Developing states argued that the transfer of resources by developed countries was not a charitable duty but a necessary obligation arising from its disproportionate use of the atmosphere.[109] They, therefore, emphasised that

[102] Ambuj Sagar and Tariq Banuri, "In Fairness to Current Generations: Lost Voices in the Climate Change" (1999) 27 *Energy Policy* 509–14; Jekwu Ikeme, "Equity, Environmental Justice and Sustainability: Incomplete Approaches in Climate Change Politics" (2003) 13 Global Env'l Change 200.

[103] Kevin Gray and Joyeeta Gupta, "The United Nations Climate Change Regime and Africa" in Beatrice Chaytor and Kevin Gray (eds), *International Environmental Law and Policy in Africa* (Springer 2003) 65.

[104] Daniel Bodansky, "The United Nations Framework Convention on Climate Change: A Commentary" (1993) 18 YJIL 479; Mickelson (n 97); Erik Paredis, Gert Geominne, Wouter Vanhove, *et al*, *The Concept of Ecological Debt: Its Meaning and Applicability in International Policy* (Academia Press 2007) 7.

[105] Michael Grubb, "Seeking Fair Weather: Ethics and the International Debate on Climate Change" *International Affairs* (1995) 71(3) 464.

[106] Andrew Hurrell and B Kingsbury, "An Introduction" in Andrew Hurrell and B Kingsbury (eds), *The International Politics of the Environment: Actors, Interests and Institutions* (Clarendon Press 1992) 39.

[107] Duncan French, "Developing States and International Environmental Law: The Importance of Differentiated Responsibilities" (2000) 49 Int'l & Comp L Quarterly 37.

[108] *ibid*.

[109] *ibid*.

the provision of resources to developing countries does not confer developed countries with additional rights or privileges in climate change governance. They thus sort for equal representation in the financial and institutional structures of the UN framework on climate change. At the point of adopting the UNFCCC, the Secretary-General of the UN made it clear that "the recognition that the industrialised countries should take the lead in tackling climate change is one of the political cornerstones of the Convention."[110]

It is clear from the above that for developing countries, equity in the climate change regime is based on two principal grounds: the historical responsibility of developed countries for the emission of greenhouse gasses and their higher capability to bear the cost of climate change. Historic equity relies on the fact that developed countries are responsible for the bulk of past and continuing GHG emissions which have led to climate change. The argument, therefore, goes that it will be inequitable to require developing states, who are mostly 'innocent' of this vice and worse still, lacking the capacity to bear equal responsibilities with developed states in the attempt to reduce greenhouse gas emissions.

It brings the application of equity in climate change law within the historical context of the activities that have led to climate change. Springer illustrates this drawing from Locke's Second Treatise on Civil Government. He explains that all humanity owns the earth and all that is in it in common. Thus, let us imagine that there was a giant sink into which we could all put our waste, everyone would be entitled to put their waste in there for as long as is necessary. Even if some put in more than others, this would create no problem where the sink has a limitless capacity to receive those wastes. However, where this sink has a limited capacity and is filling up already, those who use it disproportionately more than others, deny the others the opportunity to use it.[111] The wrongful arrogation of this sink in the past by some group of people, therefore, calls for compensation of those who now have to use the sink restrictedly despite their non-contribution to the disproportionate use in the past.[112] The position is therefore that for any distribution in the present to be equitable, it must take this history into account in its distribution of costs and benefits.

Developing states, therefore, emphasises the role that developed states have played in emitting greenhouse gases and on that basis, developed countries should bear the necessary costs of climate change mitigation. Past greenhouse gas emissions should be considered in deciding who bears the cost of climate change mitigation

[110] Kofi Annan, "We need decisive action in Kyoto to limit Greenhouse Gas Emissions" (1997) 15(4) *Climate Change Bulletin* 4.

[111] Peter Singer, *One World: The Ethics of Globalization* (3rd Ed, Yale University Press 2016) 32.

[112] *ibid* 37.

and adaptation and in the distribution of emission cuts and targets.[113] Bygones are not bygones, rather historical emissions are to be taken into consideration in sharing the cost of climate change, and since "developed countries are responsible for the present damage, then it is their responsibility to clean the mess."[114] It thus invokes historical principles of justice which demands that a person should bear responsibility for the outcome of their activities.[115]

This not only applies to the question of who should bear the cost of climate change but also to the allocation of emission targets to different countries. The earth and its ability to absorb emissions are seen as a shared resource, a question of justice, therefore, arises as to how this should be distributed among various countries. In the story of the sink, it is like deciding who should get how much of sink space given the disproportionate use in the past. In order to be equitable, recourse must necessarily be made to how much of this resource each country has explored in the past before giving allocations. Reducing emissions would mean a reduction in the general level of emissions and certain industrial activities would either need to be stopped or restructured since increased greenhouse gas emissions have traditionally been a by-product of industrialisation. Developing countries, therefore, argued that it would be inequitable to undermine their development through emission targets and argued that they should be allowed emissions reaching the same per capita level with the emissions of developed countries.[116]

Developing countries have also argued that developed countries should take the lead on climate change based on their higher levels of wealth and capability. This position relying on the first holds that since developed countries enjoy higher levels of wealth and technological development, they are in a better position to bear the cost of climate change. Equity here thus focuses on capability rather than culpability. This is based upon the ideas of fairness that the condition of those who are worse off should be improved by those who have the means to do so.[117] It gives the responsibility to the one who can bear it the most. These arguments focus on the need for redistributive justice. The redistribution of wealth by developed

[113] Mathias Friman, *Historical Responsibility in the UNFCC* (Centre for Climate Science & Policy Research 2007) 2.

[114] Ikeme (n 102) 201.

[115] Stephen M Gardiner, "Ethics and Global Climate Change" (2004) 114(3) *Ethics* 570. It is worthy of note that developing countries successfully got the rest of the world to acknowledge the need to restructure the Global Environmental Facility and make its operation more democratic, transparent and financed as appropriate. Articles 11 and 21(3) of the UNFCCC.

[116] Jonathan Wiener, "On the Political Economy of Global Environmental Regulation" (1999) 87 Georgetown L.J 765; Patricia Birnie, Alan Boyle and Catherine Redgwell, *International Law and the Environment* (3rd Ed, OUP 2009) 374; David Pearce and Kerry Turner, *Economics of Natural Resources and the Environment* (Johns Hopkins University Press 1990) 61–9.

[117] Singer (n 111) 46.

states to the least advantaged members of the society has been described as a desirable end if it would lead to the most significant benefit.[118] Rawls has noted that this seeming inequality that places higher responsibility on certain states is not necessarily unjust.[119] He notes that if the "basic structure of the Society of Peoples" contains certain unjustified inequalities, then it would be necessary to correct such inequalities.

Although closely related, compatible and reinforcing of each other, these two grounds are distinct. Whereas historic equity sees the earth as a common resource of humanity and seeks to hold developed states accountable for climate change damage, the second ground focuses more on the ability of respective countries and the principle of distributive justice rather than the cause of climate change. The collection of countries illustrates this difference in Annex 1 under the Convention. There were no criteria for deciding which country belonged to the group. For example, a country such as Switzerland who is generally not guilty of historical emissions is also included in Annex 1.

Developed countries are therefore historically responsible for over seventy per cent of historical global emissions of GHG while developing countries contribute about a quarter even though they have eighty per cent of the world's population.[120] As at the early 2000s, the United States alone was responsible for almost twenty-five per cent of the world's greenhouse gas emissions as compared to 136 developing countries who contribute a total of twenty-four per cent of these emissions.[121]

The question at the time of concluding the UNFCCC was therefore whether developing countries who, at that time, had played little or no role in the historic emissions leading to climate change should be saddled with the same responsibility as developed countries, whose actions had been the major cause of climate change. The contentions at the Rio conference, therefore, centred around who should take responsibility, what that responsibility should be and under what circumstances should these responsibilities be discharged. The principle of equity and common but differentiated responsibilities was therefore introduced into the UNFCCC as a rallying point for the contention that climate change was a universal responsibility of all mankind on the one hand and the contention that certain countries should bear the cost of addressing climate change having been largely responsible for historical emissions of GHG.[122] The underlying principle for equity (and CBDR)

[118] John Rawls, *A Theory of Justice* (Harvard University Press 1971) 266.
[119] *ibid* 113–4.
[120] Olivier (n 99) 34; Ashton (n 99) 5.
[121] Roberts (n 100) 393.
[122] This twin principle has since occupied a central position in climate change law and policy.

under the climate change regime was the same with earlier introduction of the term in both domestic and international law: the need to do fairness and justice. In this case, the need to be fair and just in the distribution of the cost of climate change given the disparity in the contributions to climate change and the different levels of emissions.

A. Equity and CBDR in the UNFCCC

The principle of equity stands at the core of the UNFCCC. Article 3 provides that parties shall be guided by the principle of "equity and in accordance with their common but differentiated responsibilities and respective capabilities." This means that even though all states have a shared responsibility for addressing the state of the climate, countries would take up different levels of obligations in tackling the problem of climate change and each country's capacity would determine this. The principle recognises the peculiar needs of the least developed states and emphasises that even though all states have a joint responsibility to protect the environment, the contribution to the cost of climate change action would differ.

It represents an acknowledgement of the fact that states differ in their ability to respond to environmental challenges.[123] While developed countries have the technological and financial capacities, developing countries have fewer of these necessary tools for responding to climate change.[124] The principle of CBDR, therefore, takes cognisance of these different levels of capabilities and states that this shall determine the extent of each state party's rights and obligations under the climate change regime. The obligations to be undertaken by each country was to be different and based on each country's national circumstance.

In outlining specific commitments to be undertaken by country parties, Article 4 of the UNFCCC places the cost of climate change action on developed countries by obliging them to provide necessary finance and technology for mitigation and adaption to climate change.[125] It is also important to note that the UNFCC places no obligation on developing countries to undertake emission cuts, while developed countries are required to commit to emission cuts. Although the Convention creates reporting obligations which bind both developed and developing states, the discharge of this responsibility by developing countries is dependent on the effective implementation by developed country Parties of their commitments to provide financial and technological resources and assistance

[123] Estherine Lisinge-Fotabong, *et al*, "Climate Diplomacy in Africa" (Climate Diplomacy Policy Brief 2017).
[124] *ibid*.
[125] Article 4(3) of the UNFCCC.

to developing countries. The Convention further provides that the discharge of responsibilities under the Convention will take fully into account that economic development and poverty eradication are the first and overriding priorities of the developing country Parties.[126]

The UNFCCC establishes a provision which conditions the implementation of obligations by developing states upon the "effective implementation by developed country Parties of their commitments… related to financial resources and transfer of technology…" This means that a developing state is not under any obligation to comply with any provision of the convention unless and to the extent to which it receives technological and financial assistance under the convention. Secondly, no commitment to climate change action must limit the ability of a developing country party to pursue or attain economic development.

The UNFCCC thus adopts a view of equity which ensures absolute differentiation between developed (Annex I) and developing countries (otherwise known as non-Annex 1 countries). Developed countries are obliged under the Convention to bear the cost for climate change action. The application of the law in this manner was necessitated by the economic and developmental needs of developing countries. It has been noted that this level of differentiation increased or motivated the participation of developing states in the emerging climate change regime.[127] It creates incentives for them to be a part of the regime by providing for differentiated standards and the transfer of finance and technology for taking climate action.[128] Scholars have noted that further consensus on higher standards in international climate change law would have been impossible to achieve without this application of equity.[129] Thus, the UNFCCC provides for an equitable principle which broadly groups countries into developed and developing countries, expressly placing obligations on the former.

The principle of equity within the climate change regime must, therefore, be understood in the light of this. The distribution of duties and responsibilities is to be dependent on each country's national circumstance. The pre-eminence of the national situation is inherent in the application of equity in the UNFCCC. Equity was not to apply blindly or permanently. A country takes up responsibilities under the regime based on its economic groupings (the UNFCCC defines state parties into Annex I, II and non-Annex I parties). Although not expressly mentioned in the UNFCCC, it logically follows that the specific roles of a state would change as

[126] Article 3(4) of the UNFCCC.

[127] Anita Halvorssen, *Equality among Unequals in International Environmental Law Differential Treatment for Developing Countries* (Westview Press 1999) 3–4.

[128] Philippe Sands and Jacqueline Peel, *Principles of International Environmental Law* (3rd Ed, CUP 2012) 43.

[129] Birnie (n 116) 135.

its economic situations and groupings also change. This implies that countries with no level of responsibility when the UNFCCC was adopted in 1992 would take on increasing levels of responsibilities as their economic situation changes.

Although (former) developing state parties may be opposed to this interpretation, it must be noted that although the UNFCCC recognises the historic emissions of developed countries, the principle of equity was not to apply based on historic emissions but based on present "respective capabilities".[130] It would, therefore, be contrary to the provisions of the UNFCCC for a newly industrialised state party to claim that it would not vary its level of responsibility to reflect its new developmental status because its historic emissions were minimal. The legal provision of the UNFCCC distributes responsibilities based on capabilities and not historical emissions.

B. Equity and CBDR in the Kyoto Protocol

The Kyoto Protocol to the Climate Change Convention was adopted in December 1997.[131] This Protocol had come into force amidst contentions between state parties that there was a need to strengthen the commitments of developed state parties under the framework convention.[132] The aim of the Kyoto Protocol was, therefore, to set specific objectives for emission reductions with specified timeframes. It was an instrument to flesh out in specific terms the general principles that had been provided for in the UNFCCC and was targeted primarily at creating specific obligations for developed countries, developing state parties were not to take up any new obligations.[133]

The process that culminated into the Kyoto Protocol has however been described as one of "the most difficult and complex ever conducted for a multilateral environmental agreement."[134] The contentions had revolved around issues such as emission reduction targets, emissions trading, joint implementation, financing, transfer of technology, the role and treatment of developing countries:[135] all issues with implications of equity within the Protocol. The divisions became heavily contested that the United States, despite being one of the major negotiators of the

130 Article 3 of the UNFCCC.

131 Kyoto, 10 December 1997, 16 February 2005; reprinted at 37 ILM 22 (1998).

132 Decision 1/CP.3, Report of the Conference of the Parties on Its Third Session, Kyoto, 1–11 December 1997, FCCC/CP/1997/7/Add.1.

133 Decision 1/CP.1, Report of the Conference of the Parties on Its First Session, Berlin, 28 March–7 April 1995, FCCC/; Preamble to Article 10 of the Kyoto Protocol. CP/1995/7/Add.1, para. 2(b)

134 Sands (n 128) 284.

135 *ibid*.

Protocol and the world's largest emitter, refused to ratify the Protocol. Nevertheless, the Protocol entered into force in February 2005.

The outcome of the Kyoto Protocol has been described as "narrow, thin and ultimately symbolic."[136] The failure of the Kyoto Protocol has been alluded to issues such as the refusal of emerging developing states to undertake commitments, the bindingness of emission targets, the withdrawal of significant emitters from the Protocol, among others. As the end of the first commitment period of the Kyoto Protocol began to approach, state parties entered into negotiations towards a second commitment period. Unfortunately, it became difficult to reach an agreement. Countries like Japan, Russia and Canada had stated that they were not prepared to be a part of the second commitment period.[137] Canada ultimately withdrew from the Kyoto Protocol in 2011.

Given the issues surrounding the 'fall' of the Kyoto Protocol, important questions in the climate change regime became how should a post-Kyoto Climate agreement deal with the emissions of and commitments of developing states, would a legally binding agreement be more desirable or should the agreement be in form of a soft law? Parties ultimately agreed at the Durban Conference of the Parties 2011 to "develop a protocol, another legal instrument or an agreed outcome with legal force under the Convention applicable to all Parties."

C. Equity at COP21

In 2011, about two decades after the UNFCCC was adopted, the Ad Hoc Working Group on the Durban Platform for Enhanced Action (ADP) was formed. This group was responsible for the formulation of a new legally binding agreement for the regulation of climate change action.[138] The Paris Agreement was adopted four years later at the Conference of the Parties (COP) held in Paris, France in December 2015. Like the climate change agreements before it, the ADP and negotiating parties had to battle with the distribution of the costs of climate change action. It is therefore not surprising that twenty-five years after the UNFCCC, equity still stood at the centre of the Paris Agreement.

The inclusion of equity under the Paris Agreement reflects the need to continue the 'burden-sharing formula' established under the UNFCCC whether as a necessity for justice under the climate regime or as a practical means for enhancing

[136] Robert Keohane and David Victor, "The Regime Complex for Climate Change" (2011) 9(1) Perspectives on Politics 5.

[137] D Ryan, *et al*, "Climate Change after Cancun: A Post-COP-16 Analysis" (2010) 18(6) Environmental Liability 208.

[138] Concern for a legally binding agreement had arisen due to the lack of compliance with the emissions reduction targets of the Kyoto Protocol. Parties, therefore, sought for an agreement which state parties will be obliged to comply with.

ambition. However, while parties generally expressed support for the continued inclusion of equity and CBDR within the climate regime, there were opposing views on the practical implications of this on the obligation of parties.[139] This was primarily because unlike the case in 1992, several 'developing' countries had now become major economies and had also become significant emitters. Countries like Brazil, South Africa, Mexico, China and India had become responsible for a considerable percentage of global emissions.

The concerns, therefore, centred around firstly, the fact that the capacities of certain countries had changed and their level of responsibility under the climate change regime should change to reflect this new capability. Secondly, developing countries had joined the industrial revolution, and there were concerns about the levels of their emissions and the need to place restrictions on these emissions. The negotiators were thus faced with the best way to reflect the principle of equity while taking these issues into consideration. The success of COP21 was to be determined, among other things, by the ability to reach a common understanding of the concept of equity and differentiation and the practical implications of this in the new Agreement.[140]

D. Equity and differentiation under the paris agreement

In the Paris Agreement, equity and CBDR took on a new form from what had been known in previous climate change agreements. Article 2 of the Agreement provides that: "This Agreement will be implemented to reflect equity and the principle of common but differentiated responsibilities and respective capabilities, *in the light of different national circumstances*" (emphasis added).[141]

The Agreement thus includes the phrase "in the light of different national circumstances", which has been described as a "dynamic approach" introduced by the Paris Agreement in the application of equity and differentiation.[142] This phrase was introduced in an attempt to ensure an adaptation of equity and differentiation, which suited the various groups and parties involved in the negotiation process. Scholars like Rajamani has therefore described equity as espoused in the Paris Agreement as distinct from equity as used in the UNFCCC. She attributes this to

[139] International Institute for Sustainable Development (IISD), "Summary of the Paris Climate Change Conference: November–13 December 2015" Earth Negotiations Bulletin 12.

[140] Christina Voigt, "Equity in the 2015 Climate Agreement: Lessons from Differential Treatment in Multilateral Environmental Agreements" (2014) 4 Climate Law 51.

[141] Article 2 of the Paris Agreement; Lavanya Rajamani, "Ambition and Differentiation in the 2015 Paris Agreement: Interpretative Possibilities and Underlying Politics" (2016) 65 Int'l & Comp L Quarterly 506.

[142] Sandrine Maljean-Dubois, "The Paris Agreement: A New Step in the Gradual Evolution of Differential Treatment in the Climate Regime?" (2016) 25(2) Reciel 153.

the addition of the qualifier "in the light of national circumstances." This addition reflects an operationalisation of equity.[143] In other words, it tailors the applicability of equity to specific issues, thus ensuring different forms of differentiation for different parties and in different areas.[144] Ultimately, the responsibilities of states will evolve as their national circumstances also evolve.[145]

Maljean-Dubois explains this caveat as a specification within the Agreement that equity and differentiation were evolutionary. She points out further that this clause gives room for a "flexible and evolutionary interpretation of the CBDRRC" as the national political, social and economic circumstances of state parties change.[146] According to her, this flexibility is the hallmark of equity and differentiation under the Paris Agreement, compared to earlier climate change agreements.[147] It serves as an insurance for developed state parties who had been concerned from the days of the Kyoto Protocol about the role of emerging economies in climate change. Developing countries will thus be obliged to move progressively towards economy-wide emissions reduction targets and to contribute to the global climate finance funds.

Bodansky also joined in the appraisal of the caveat.[148] He sees it as introducing an element into the climate change regime, which ensures that equity and differentiation do not exist in a vacuum but in accordance with the different state of each country.[149] Another author describes it as broadening the range of factors on 'the basis of which Parties can be differentiated beyond responsibility and capacity' and thus shifting from the bilateral approach which had been dominant in the preceding instruments.[150] According to Doelle, it is a more nuanced approach to differentiation.[151] All of these scholars hinge this appraisal on the same reasoning: that the phrase 'in the light of different national circumstances'

[143] Rajamani (n 141) 509.
[144] *ibid* 508–9.
[145] *ibid* 508.
[146] Maljean-Dubois (n 142) 153–4.
[147] *ibid* 154.
[148] Daniel Bodansky, "The Paris Climate Change Agreement: A New Hope?" (2016) 110 Am J Int'l L 288–319.
[149] *ibid* 299–300.
[150] Meinhard Doelle, "The Paris Agreement: Historic Breakthrough or High Stakes Experiment?" (2016) 6 Climate Law 15.
[151] *ibid*.

the principle of equity and differentiation is applied in line with the prevailing or current conditions and circumstances of a state party.

VI. In the light of different national circumstances: more words, less meaning[152]

Despite the position of these scholars, a clear reading of the principle of equity and differentiation as contained in the UNFCCC reveals that the phrase, "in the light of different national circumstances," adds nothing new to the principle. It is necessary to understand the nature of the principle of equity and differentiation and to understand the kind of treatment it offers developing state parties. This principle was introduced in the UNFCC based on two overarching issues: the responsibility of developed countries for historical emissions and the low capacity of developing states to respond to climate change.[153] Thus, the responsibility of states is made to correlate with their level of capacity.

In other words, state parties enjoy a contextual treatment. Magraw defines contextual treatment as a norm which requires that different factors be taken into consideration with respect to different countries. [154] In other words, certain circumstances (context) determine the kinds of and the level of obligations to be borne by each country. For equity and differentiation, this circumstance is the '*respective capabilities*' of a country to respond to climate change. In other words, the level of obligation of a state party under the UNFCCC will be determined by its capacity to undertake climate action. Since a country's economic condition will usually not be static, it implies that its obligations under the UNFCC climate change regime will also not be static. Levels of obligation will change as the economic realities and abilities of the state changes.

This is also illustrated in the specific phrasing of the principle; it reads "common but differentiated responsibilities and *respective capabilities*." The term 'respective capabilities' means that states have different capacities and the specific obligations to be undertaken would reflect this difference in capacity. In other words, capacity determines obligation. The principle acknowledges that the circumstances and hence capacities of each country would differ, and this difference would determine their responsibilities. Where a country that hitherto lacked capacity improves in its economic fortunes, it follows that its 'respective capability' now

[152] The phrase 'in the light of different national circumstances' had previously appeared in the US-China Joint Announcement on Climate Change (11 November 2014).

[153] Preamble to the UNFCCC

[154] Daniel Magraw, "Legal Treatment of Developing Countries: Differential, Contextual and Absolute Norms" (1990) 1 Col J Env'l L & Policy 1.

allows it to be able to take on more commitments and its obligations would change to reflect the change in its current capabilities.

Thus, the addition of the phrase "in the light of different national circumstances" in the Paris Agreement merely states the obvious. It repeats a statement which is inherent in the principle of equity and differentiation as couched in the UNFCCC. It adds nothing new to the application of equity and differentiation in global climate change governance. The underlying principle of equity and differentiation is that states are treated based on their different capabilities and the phrase merely restates this. The only other usefulness of the addition may be found in whether it helps clarify the principle of equity and differentiation and how it should apply. Unfortunately, it does not. It gives no clarity on how a change in circumstance should be determined, to what extent must the economic capabilities of a state party change before its obligations under the treaty change, to what extent should a change in capabilities correlate with the level of change in the level of commitments, *etc*.

Having said that, it is necessary to point out that the Paris Agreement indeed introduces a new dimension to the application of equity and differentiation under the global climate change regime. However, these are to be found in other parts of the Agreement and not in the addition of the phrase "in the light of different national circumstances." A major shift in the Agreement is the absence of a list of 'annexed' countries as seen in the UNFCCC and the Kyoto Protocol and the provision on climate funds. These changes help to reflect the changing *capabilities* of country parties. As national circumstances have changed the grouping of parties into Annexes was no longer relevant, and the notion that developed countries be the sole donors of climate finance gave room for a provision which encouraged others to contribute to such funds. The Paris Agreement also contains other concepts such as sustainable development,[155] equitable access to development,[156] poverty eradication[157] and climate justice[158] which are closely related to and reflect the principle of equity. We shall take a closer look at the application of equity and differentiation to climate finance under the Agreement.

VII. Climate finance and equity under the paris agreement

The issue of climate finance is one of the significant areas of the climate change regime, where the application of equity and common but differentiated responsibilities and respective capabilities (CBDRRC) has enormous implications.

[155] Preamble to the Paris Agreement, articles 2, 4, 6, 7, 8 and 10(5).
[156] Preamble to the Paris Agreement.
[157] Preamble to the Paris Agreement, articles 3, 4 and 6(8).
[158] Preamble to the Paris Agreement.

Under the UNFCCC, financial obligation was bifurcated, developed countries being the donors while developing countries were the beneficiaries. This was in line with equity and differentiation since, at the time, developing countries generally lacked the capacity to fund climate action. Although the Paris Agreement adopts the traditional position that "developed country Parties shall provide financial resources to assist developing country parties with respect to" climate action and obliges them to take the lead in the mobilisation of climate finance in favour of developing countries,[159] it introduces certain shifts from the traditional UNFCCC position on climate finance. Also, these shifts are very significant in understanding the application of equity and differentiation within the Paris Agreement, and international climate change law in general.

Article 9(2) of the Agreement provides that "other Parties are encouraged to provide or continue to provide such (financial) support voluntarily."[160] In other words, while not making it obligatory, the Agreement seeks to widen the class of donor states by including donors other than developed country parties, albeit voluntarily. The inclusion of this phrase had followed the desire by countries such as the US, Canada and Australia to broaden the category of donor states by introducing a new group with a higher level of commitment and to be known as 'graduated parties', 'major emitters' or 'advanced developing countries'.[161] According to them, this would be a proper reflection of the present economic realities and financial capacities of different states.[162] This was, however, rejected by the majority of the developing countries to be affected by this new categorisation.[163] Ultimately the term 'other parties' was adopted, and these other parties were merely 'encouraged' to contribute, and no obligation was created.

Thus, the Paris Agreement while maintaining the traditional UNFCCC arrangement wherein developed state parties have the primary obligation to provide climate finance, creates a new category of 'other parties' who are encouraged to contribute towards climate finance. This reflects the change in the economic and developmental level of some developing countries such as Brazil, China, India, Mexico, Singapore and South Africa. The position was thus held that since there had been a change in their level of development since the UNFCCC

[159] Article 9(3) of the Paris Agreement. Parties also set a target of $100 billion per annum in climate finance till 2025. Unfortunately, this target has been largely unmet.

[160] Article 9(2) of the Paris Agreement.

[161] Sikina Jinnah, *et al*, "Tripping points: barriers and bargaining chips on the road to Copenhagen" (2009) 4 Env'l Res Letters 3.

[162] Ralph Bodle, Lena Donat and Matthias Duwe, "The Paris Agreement: Analysis, Assessment and Outlook" (2016) CCLR 11.

[163] Rajamani (n 144) 508.

was first adopted, equity must be adopted in a manner that reflects this change in circumstance.

In conclusion, equity and differentiation remain at the centre of global climate change law and policy. The implication of this principle on the specific commitment of each country is, however, not static. It would evolve as the economic situation and general capabilities of a country changes. This is reflected in the very wording of the UNFCCC which provides for 'respective capabilities.' The Paris Agreement has helped flesh this out in various areas of commitments, the most obvious of which is the provision of climate finance. This changing level of obligation is the very core of differentiation according to respective capabilities. The addition of the phrase, "in the light of different national circumstances," adds nothing to this principle or its application within the climate change regime. The Paris Agreement's contribution to the jurisprudence and application of equity and differentiation is to be found in its various provisions on mitigation, adaptation, the absence of a list of annex countries and importantly, its provision on climate finance.

VIII. Conclusion

The history of equity and CBDR in the global climate change regime is a long one. It is a history embedded in notions of justice and the need to ensure that the costs and benefits of climate change are equitably distributed. This means that those who have contributed the least to climate change will not be made to bear the cost of climate change in a manner that is disproportional to their contribution to global emissions and their capacities. This was not a problematic categorisation to make at the inception of the climate change regime, and the UNFCCC clearly delineates country parties into different groups known as Annexes.

However, twenty-five years after the UNFCCC, developing countries have emerged into bigger economies and industrial nations. Some of them now contribute immensely to climate change and have developed financial capabilities such that obliging them to take climate action will not impede their development. This has necessitated the need to ensure that the new Agreement reflects the changed capacity of certain developing country parties. Although the Agreement adds a caveat "in the light of different national circumstances" to the phrasing of the principle of equity and CBDR under the Paris Agreement, this addition does nothing in fleshing out the practical implication of differentiation on specific commitments of country parties. The changing capacities of state parties are reflected in other parts of the Agreement and more especially Article 9 on climate finance, with its subtle provision for a new class of climate fund donors. The

Agreement also does away with the old method of annexing countries as seen in the UNFCCC and Kyoto Protocol.

The Agreement thus ensures that a generalisation of countries into developed and developing countries irrespective of their *respective capabilities* is avoided. This is achieved not with the additional phrase to the provision of equity and CBDRRC but through the nuance which the Agreement introduces into its various substantive provisions on commitments and obligations. The novelty of the Agreement is thus to be found within its other provisions. These include the abolition of the annex of state parties, sustainable development, and equitable access to development, poverty eradication and climate justice.

Cambridge Law Review (2019) Vol IV, Issue ii, 138–147

Sumangali System: Is the Truth Ugly?

Yamuna Menon*

Abstract

The tussle between universalism and relativism is a hotly debated issue under the human rights regime. The influence of the same is visible in varied spheres including labour rights. India in particular being home to a huge labour force demands attention on the various local practices in the industries and its contrast with rights guaranteed to labour at a universal level. The black letter law does not always produce the ideal solution, and one such instance is the Sumangali system in the textile industry. However, is this system extremely evil as it is presented to the world? The answer requires a deeper and closer look at the context and the lives of the stakeholders.

I. Introduction

The Universal Declaration of Human Rights emphatically states that "[e]veryone is entitled to all the rights and freedoms… without distinction of any kind" under Article 2. This is the core spirit of universalism, wherein the rights must be enjoyed by individuals based on universal notions.[1] The labour standards set by International Labour Organisation for instance, stems from such ideas that are placed as standards for the nation state.[2] However, has this idea then triumphed

* Menon is a final year B.A. LL.B. (Hons.) student at National Law School of India University, Bangalore. I would like to acknowledge and extend my heartfelt gratitude to Prof. Babu Mathew and Prof. Rashmi Venkatesan for providing helpful insights and sharing their expertise.

1 J Donnelly, "Human Rights: Both Universal and Relative (A Reply to Michael Goodhart)" (2008) 30(1) Human Rights Quarterly 194, 195.

2 International Labour Organisation (ILO), "How International Labour Standards are Created" <https://www.ilo.org/global/standards/introduction-to-international-labour-standards/international-labour-standards-creation/lang--en/index.htm> (accessed 15 April 2019).

in extending recognised universal rights and freedoms to every human across jurisdiction? The response would be a clear negative. It is in this context scholars suggested economic relativism as an explanation to differential rights and freedoms since these are founded on socially constructed ideas that differ according to time, place, nation and culture.[3] This leads to a friction between universalism and economic relativism in many instances. The widely prevalent Sumangali system in the textile industry of Tamil Nadu is a suitable example in order to understand these conflicting principles on human rights discourse.

The main focus of this paper is to understand the reasons behind the continued prevalence of Sumangali system that blatantly infringes human dignity. Why does this system attract girls? Why has this system attained the significance it has today? Is it a trade-off between rights and 'development'? Should this system continue? The paper will delve into these in light of economic relativism. With these questions in the backdrop, the central argument of the paper is that *Sumangali system thrives on the 'fictional' female agency of the girl child*.

The theme is developed around the contrast between universal labour rights and economic relativism analogous to cultural relativism, by relying on academic articles, reports and books. The aim of the project is to bring out the core reasons for the flourish of the Sumangali system in textile industry, and the objective is to analyse the extent of agency exercised. For this, the real-life experiences of the young girl children provided the perspective to pierce the veil of dominant narrative of 'victimisation'.

II. What is sumangali system?

Sumangali scheme[4] is a system prevalent in the Tamil Nadu textile industry, wherein girls aged between 13–18 years are employed as 'apprentice'.[5] The scheme witnesses migrant workers moving in to these textile mills from the neighbouring states as well.[6] The idea of Sumangali worker is that of a female

[3] D Donoho, 'Relativism versus Universalism in Human Rights: The Search for Meaningful Standards" (1991) 27(2) Stan J Int'l L 345, 348.

[4] Exchange of dowry is a recognised crime under Dowry Prohibition Act, 1961.

[5] Sumangali scheme and bonded labour in India, Fair Wear Foundation (2010) <https://www.fairwear.org/wp-content/uploads/2016/06/fwf-india-sumangalischeme.pdf> (accessed 15 April 2019). As per a later decision in *Tamil Nadu Spinning Mills v The State of Tamil Nadu* Working Paper No. 9182 of 2007, apprentices are covered under Minimum Wages Act, 1948. Hence non-payment of minimum wages to Sumangali girls is violation of law.

[6] Gerard Oonk, *et al*, "Maid in India: Young Dalit Women Continue to Suffer Exploitative Conditions in India's Garment Industry" (Report, Centre for Research on Multinational Corporations and India Committee of the Netherlands, April 2012) 19; *Inter-State Migrant* Workmen *Act, 1979 has been enacted in India for governing matters of migrant workers.*

who works in these mills for a collated amount after a period of time to facilitate marital bliss. In essence, the narrative accepted by many is that the girls work for earning their potential dowry and the system is a marriage assistance scheme.[7] Moreover, the formal employment relationships are avoided in order to escape the obligations imposed on employers at multiple levels.[8] Therefore, this system is considered to be exploitative in nature where the innocent girls are termed to be the 'victims' of the greed of capitalism.[9] However, many fail to notice the other side of the story; the story that the Sumangali girls have to say. The civil society assumes upon themselves the moral obligation to free the victims from the shackles of Sumangali scheme.[10] But is that what the girls want? The proponents of rights and freedom simply cut through the system, without understanding the core needs leading to the flourish of such system in the twenty-first century. The fundamental reason for this biased view 'to protect the girls' arise from the idea of universalism and the uniform application of labour standards at the international level to the local labour force.

A. Universal labour rights

ILO conventions and the sustainable development goals[11] suggest certain standards of labour rights that the nation states must strive to achieve.[12] India has been falling behind in taking up this responsibility until last year when it ratified

[7] Sindhu Menon, 'Adolescent Dreams Shattered in the Lure of Marriage: Sumangali System: A New Form of Bondage in Tamil Nadu' (2006) 4(3) Labour File <http://www.labourfile.com/section-detail.php?aid=337#> (accessed 15 April 2019). Additionally, during my conversation with NGOs in Tamil Nadu (CARET), the idea followed by civil society is that of a direct connection between Sumangali system and dowry.

[8] Apprenticeship Rules Central, April 2015. Apprentices cannot be more than 10% of the workforce which is violated by the textile mill owners. They employ girls under the Sumangali system without formal contracts.

[9] N Mani and N Krishnan, "Understand the Labourer's Problems Under the Sumangali Thittam Scheme in Textile Industry in Tamil Nadu, India" (2014) 1(6) International Journal of Business and Administration Research Review 118, 120.

[10] Understanding the Characteristics of the Sumangali Scheme in Tamil Nadu Textile & Garment Industry and Supply Chain Linkages, Solidaridad and Fair Labor Association (2012) 27.

[11] Target 8.7 which suggests to "take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms" under Sustainable Development Goal 8-Target 8.7 (Promote Inclusive and Sustainable Economic Growth, Employment and Decent Work For All), International Labour Organisation.

[12] Also see Universal Declaration of Human Rights 1948, articles 23 and 24.

two conventions concerning child labour.[13] Subsequent to this, policies were to be formulated to implement the ideas contained in these conventions. But can we implement these ideas applicable on a universal level at equal par in India? Here is where the idea of universalism stumble.

Universal labour rights envisage elimination of child labour so that children are facilitated with the opportunity to develop physically and mentally to their full potential.[14] In addition to this, the Minimum Age Convention, 1973 provides that:

[M]inimum age for admission to any type of employment or work which by its nature or the circumstances in which it is carried out is likely to jeopardize the health, safety or morals of young persons shall not be less than 18 years.[15]

Economic exploitation of a child is looked down upon by the international community.[16] However, the context and circumstances of the vocal nations while framing such obligations at international level differ considerably from developing nations like India, for instance.

The scholars supporting universalism would consider that, for example, concepts like equal remuneration,[17] prohibition on forced labour,[18] safe housing conditions and facilities[19] is universally applicable and all nations must adopt steps to ensure the same. In furtherance, India has made efforts to follow certain international obligations.[20] However, still our society has not escaped the vices of exploitation of young labour and has not been successful in providing dignified working conditions. This is because of a crucial reason—the idea of rights and freedom differ according to place, time and culture.[21] The way an American labourer asserts her right to clean workspace is to be distinguished from an Indian labourer in content as well as context in its entirety.

Economic relativism comes in handy in such situations when the idea of one's right has to be shaped based on the circumstances and surroundings.

[13] Minimum Age Convention 1973 (No. 138) and the Worst Forms of Child Labour Convention 1999 (No. 182) were recently ratified by India.

[14] International Labour Office (Geneva), *The End of Child Labour: Within Reach* (International Labour Conference, 95th Session 2006) <https://www.ilo.org/public/english/standards/relm/ilc/ilc95/pdf/rep-i-b.pdf> (accessed 15 April 2019) 23–9.

[15] Minimum Age Convention, ILO C 138, article 3. Read also, Worst Forms of Child Labour Convention 1999.

[16] United Nations Convention on the Rights of the Child 1989.

[17] Equal Remuneration Convention 1951, ILO C 100.

[18] Elimination of Forced Labour Convention 1957, ILO C 157; Discrimination (Employment and Occupation Convention) 1958, ILO C 111.

[19] Recommendation 115 on Workers' Housing Recommendation 1961.

[20] The Child Labour (Prohibition and Regulation) Amendment Act 2016, The Right of Children to Free and Compulsory Education Act 2009 and Tamil Nadu Hostels Act 2014.

[21] Ben White, "Defining the Intolerable Child Work, Global Standards and Cultural Relativism" (1999) 6(1) Childhood 133, 134.

Universal labour rights cannot be applied on a one-size-fits-all framework, however disregarding them in the name of relativism is to be frowned upon. Economic relativism cannot in any way legitimise the violations of human rights, however it can be a practical analytical tool.[22]

B. Economic relativism as a quietist approach?

Relativism, be it economic or the larger boundary of cultural relativism, is a useful corrective tool to pseudo-universalist notions.[23] As Balakrishna Rajagopal rightly puts it "[d]evelopmentalisation of human rights discourse has caused the turn to culture".[24] Critics argue that economic relativism is quietistic in accepting the utility vis-à-vis those economies. These surroundings are the product of social hierarchies, political actions and various forms of decision making subject to moral and utility grounds.[25] Imbibing local content to global standards help in understanding and discovery and not legitimising the practices.[26] Therefore, economic relativism will be the suitable tool in understanding systems like Sumangali scheme to decipher the core reasons such practices are active in the domestic sphere.

According to economic relativism, social economy acts as the foundational 'authority' for decision making. It provides the certainty to the decision maker to justify the acts in the face of legal ambiguity, by providing a solution to the 'economic dilemma' of liberal legalism.[27] However, this justification through economic relativism at the outset seems to move away from the need to understanding the basis of practices by imposing the widely accepted standards from outside. This is a wrong approach as can be seen from the dominant narrative surrounding Sumangali system.

III. Sumangali system: issues, causes and results

Economic relativism and the development discourse have a close connection. State as a motor of economic development engages in developmental repression which is justified through the trade-off thesis.[28] When private parties engage in such repression, the idea takes a U-turn. The trajectory privileges criminal law

[22] *ibid* 135–7.

[23] Robin West, "Relativism, Objectivity, and Law" (1990) 99(6) Yale LJ 1473, 1475–8.

[24] *Balakrishna Rajagopal*, *International Law From Below: Development, Social Movements And Third World Resistance* (2003) 203.

[25] West (n 23) 1492.

[26] Donoho (n 3) 351–2.

[27] West (n 23) 1492.

[28] J Donnelly, *Universal Human Rights in Theory and Practice* (1989) 188.

approach that works via the reductive traditional dialectic of victim protection and offender incarceration. This distances itself from the direction of freedom.[29]

The imposition of state security apparatus, which are influential products of universal labour standards and rights with its discourse and technologies of control, is removed from the women's human rights, their freedom and easing their suffering.[30] Therefore, it is pertinent to understand the idea of Sumangali system as understood by the Sumangali girls themselves.

It is a matter of fact that Sumangali girls do not believe the system to be repressive in nature at the outset.[31] Is it their own belief or a belief imposed by the surroundings? It is argued that, it is the latter in light of the "fictional" agency that the girls exercise at all stages of their decision-making while being part of the Sumangali system and this unfortunately marks the flourish of Sumangali system through the lens of economic relativism.

Even though financial conditions trigger the idea to work at a young age, it is not the sole reason for girls to join the textile factories.[32] There are diverse reasons including lack of interest in education, public perception of being part of industrialisation (through a factory job) as opposed to primary agricultural hard labour and a false imagination of a luxurious life at the textile mills.[33] It was also understood that many of them make this choice themselves and not under parental pressure, subject to exceptions.[34] However, it is argued that this exercise of their female agency is distorted and is fictional in nature. This is because the reasons leading their decisions are founded upon painted circumstances that the community itself imposes on the girl child. She is not free from the chains of public opinion and perception to independently think and make choices.

These influences on the female agency are manifold. *One*, the girls who are already part of the system provide rosy picture of a future to the gullible young child—about financial independence and other 'attractive' facilities. *Two*, the factory personnel convince the girls to believe promises on money, education, protection and other services to make her life 'enjoyable'.[35] *Three*, the influence of

[29] See Ratna Kapoor, *Gender, Alterity and Human Rights: Freedom in a Fishbowl* (2018) 98.

[30] See Janie Chuang, "Rescuing Trafficking from Ideological Capture: Prostitution Reform and Anti-Trafficking Law and Policy" (2010) 158(6) University of Pennsylvania Law Review 1655.

[31] Solidaridad and Fair Labor Association (n 10) 6, 22; White (n 21) 139.

[32] Oonk (n 6) 17–8.

[33] Macro Level Understanding of Sumangali Scheme and its Impact on the Lives of Camp Coolie Workers and the Economic Share of the Camp Coolie Workers in the National And International Economy, SOCO Trust – Action Aid (2014) 5-6.

[34] Solidaridad and Fair Labor Association (n 10) 13–24.

[35] These promises are given to allure the girls into the scheme and turn out to be false later.

outside media and entertainment on the idea of her freedom and human dignity,[36] to name a few. This creates a bubble of false consciousness within which the girl child functions day in, day out.

The promise of the lump sum amount attracts majority of the girls to work under this scheme. The lack of availability of independent financial services and access to collated amounts for persons lacking security evidenced the reasoning for choosing Sumangali scheme. Marital bliss is not the sole driving force for the necessity of such funds; there are instances like medical requirements, repayment of loans and maintenance of basic shelter.[37]

When it comes to the parents of the girl child, they consider the factory/mill premises to be a safer surrounding for their girl children to grow with adequate basic facilities. Parents also consider it as a way to develop a sense of discipline and moulding of their girls to fit the traditional ideas of a role of woman in society, where the ultimate goal is to attain a happy family life.[38] The girls during their upbringing are shaped into endorsing this narrative in the society which in turn leads to their exercise of agency being biased and narrowed, "fictional" as it is called in this paper.

It is disappointing to note, while advocating the rights and freedoms of the Sumangali girls, the civil society "[w]hich is the nongovernmental and noneconomic connections and voluntary associations that anchor the communication structures of the public sphere in the society component of the life world", as Habermas calls it,[39] fails to address the root causes while in its race to highlight the exploitation and human rights violations. The diehard abolitionists who consider such practices sabotage universal labour rights[40] imposes their ideas on the girls to make them believe that their narrative is 'the' sole one in their best interest. By failing to listen to these girls, the mistake we commit is again pushing the girls to make choices under their fictional female agency.

IV. Way forward

Taking guidance from 'capability approach' as propounded by Martha Nussbaum, the Sumangali scheme must be understood in the framework of *one*,

[36] Oonk (n 6) 17–8; Solidaridad and Fair Labor Association (n 10) 6, 13–24; SOCO Trust – Action Aid (n 33) 36. It must be noted how such short term facilities or pleasures attract the girls into this scheme. This in turn shows the gullible nature of the girls who are unable to separate false from the reality.

[37] SOCO Trust – Action Aid (n 33) 36.

[38] Solidaridad and Fair Labor Association (n 10) 15, 22.

[39] Jurgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (MIT Press 1996) 366–7.

[40] Solidaridad and Fair Labor Association (n 10) 26.

primary moral importance of freedom to achieve well-being, and *two*, this freedom to be subject to the people's capabilities, their real opportunities and the value accorded.[41] Therefore, the abolitionist approach will worsen the situation by directing the girls to hardship as this view is blind to the narrative of the girls on account of over-emphasis on universalism and the universal labour rights. The societal structures and hierarchies cannot be divorced while understanding the Sumangali system as they are deeply rooted in the local context. This however, in any way, does not legitimise the violations but will provide solutions to the dilemmas over certain practices. This approach will also ensure that girls are facilitated with an environment to make choices free from other influences gradually, by better understanding the economic context, freedoms and capabilities.

The first step should be in formulating a legal regime to govern the Sumangali scheme. In doing so, it must cover specificities about the legal status of a Sumangali girl—whether she is an apprentice, contract labourer or employee. In addition, it should also contain guidelines on eligible age for work, mandatory requirement for a written contract, details on working hours, wages and leave as well as rules governing migrant workers. This express legal framework can raise awareness about the scheme and allied labour issues among various stakeholders including the international brands in the supply chain.

On the other hand, the international textile brands can also initiate steps by having tags on their clothes which mention, for instance, 'made under Sumangali' which shifts a moral burden on the ultimate consumer, who can make a choice. In the long run when the demand for such goods falls, these steps will ensure that the exploitative working systems are weeded out and replaced with healthy working environment for the girls.

Following this, there must also be efforts in developing mutual support networks with current Sumangali girls and ex-Sumangali girls. The girls who have experienced the vices of the system can share their insights as well as other opportunities available outside of the same, with girls who are currently trapped in the Sumangali system. This will help in creating an environment of mutual

[41] M Nussbaum, "Human Rights and Human Capabilities" (2007) 20 Harv Hum Rts J 21, 22. In addition to this, certain regulations governing mill owners for providing standard work environment will facilitate the achievement of this framework.

respect, understanding and empathy which can empower the girls to realise the imposed exploitation and consequently make a choice for themselves.

Lastly, the State should take steps to ensure that educational programmes, skill development sessions and other alternative avenues to explore one's potential are readily accessible in neighborhood where the Sumangali system prevails.

The imposition of views on the girls, alter their decision-making process and thereby results in only exercise of an agency, which they think are 'real', to be only a manifestation of 'fictional' female agency. This must be cut at the root for achieving long lasting solutions for the community at large by erasing information asymmetry. This will also give rise to Sumangali system not seen as a plague in the society, but as a practice which is to be understood and analysed in the local context. In conclusion, the girls have to undergo a process of cleaning the already existing notions in order to exercise their free will which will ensure dignified human life (individual human dignity) as a product of one's choice.

V. Conclusion

Universal application of labour standards to all the nations divorcing the local contexts is a relativist's nightmare. The paternalistic narrative picturising Sumangali girls as 'victims' of oppression and exploitation is lacking an understanding of the issues surrounding this practice. It is in this context that lens of economic relativism must be utilised to understand and analyse the system, and in no case to legitimise the practice.

Universal labour rights prescribe standards to be achieved and cannot serve the labour force in every nation state as they differ in place, time, culture, social hierarchies, political actions and various forms of decision making. Economic relativism on the other hand, facilitates in understanding systems like Sumangali and subsequently finds suitable solutions to the problems in such systems.

When social economy acts as the authority for decision making in cases of legal ambiguity, certain problems arise especially when the primary stakeholders in systems like Sumangali lack voice. The diehard abolitionists fail to acknowledge the reasons behind Sumangali system. It is not just a method to gain a collated amount for potential dowry; in most of the instances, the girls themselves make the choice to work for varied reasons as already discussed. However, this decision-making process is tinted with public perception, beliefs of other members of the society who directly influence the girls and other mediums of media. Therefore,

the choice that the girl child makes using her agency, is in reality not a 'real' choice, but a product of 'fictional' female agency.

Additionally, when it comes to advocating rights and freedoms of Sumangali girls, again we tend to forget that our views are imposed on the girl as opposed to letting her decide in an unopinionated environment. Therefore, it is their capabilities—opportunities and its value—that the Sumangali girls must gradually become aware of, in order to exercise their agency to full potential. As witnessed above *Sumangali system thrives on the 'fictional' female agency of the girl child*, and solely a cut through approach deciphering the reasons for decision making by Sumangali girls can ensure individual human dignity in the long run. Therefore, the truth behind Sumangali system is not as ugly as it seems to be.

Cambridge Law Review (2019) Vol IV, Issue ii, 148–175

Justice, Jurisdiction, and R (on the application of Privacy International)

Owen Alexander Sparkes*

I. Introduction

In the 2019 case of *R (on the application of Privacy International) v Investigatory Powers Tribunal and others*,[1] the Supreme Court gave a landmark judgement in which it held that the authority of *Anisminic Ltd v Foreign Compensation Committee*[2] applied to the facts of the case as an "obvious parallel". [3] At face value, the decision has simplified the law on ouster clauses. Ouster clauses are provisions that remove the courts of their judicial supervision over the decisions of executive bodies. The legislature creates these by introducing a clause into legislation that prevents judicial review of specialist tribunals' adjudicatory decisions. By their nature, therefore, ouster clauses are a controversial subject. Writing shortly after Lord Denning MR held an ouster clause invalid in the case of *Pearlman v Keepers and Governors of Harrow School*,[4] Garner announced that the "wonderful octogenarian… the Lord Mansfield of our century—has done it again!".[5]

This support for judicial intervention in the ouster clause, however, has been far from ubiquitous. Underlying ouster clauses is a debate over what governmental behaviour is permissible by the courts and what court intervention is permissible by the government: a line that divides orthodox views on parliamentary sovereignty

* MPhil (Candidate), St John's College, University of Cambridge, BA (Hons.) (Dunelm) and GDL (London). Sir Thomas More Scholar, the Honourable Society of Lincoln's Inn.

1 *R (on the application of Privacy International) v Investigatory Powers Tribunal* [2019] UKSC 22.

2 *Anisminic Ltd v Foreign Compensation Committee* [1969] 2 WLR 163 (HL).

3 *Privacy International* (SC) (n 1) [2] (Lord Carnwath).

4 *Pearlman v Keepers and Governors of Harrow School* [1978] 3 WLR 736 (CA).

5 JF Garner, "Exclusion Clauses" (1979) 42 MLR 578, 578.

and the rule of law. If Parliament is sovereign, then it can do anything, including remove the authority of the courts to review. This may permit executive bodies to act without concern that the courts will intervene if they perceive that a decision is reviewable. However, judges—aware of their constitutional role underpinned by the rule of law—are apprehensive to fully exempt governmental bodies from judicial review. As Leyland and Anthony noted, this would be "tantamount to opening the door to potentially dictatorial power".[6] Therefore, these pieces of statutory aversion to review are often met with canny judicial intervention, which is used to prevent ouster clauses precluding the justiciability of tribunals and lower courts. The decision in *Privacy International* has brought to the fore questions about judicial intervention in political bodies and is an example of how the concept of what is justiciable has changed in recent years. This paper will demonstrate how such a change is indicative in the Supreme Court's decision and reveals an incremental creep towards jurisdiction of what previously would have been considered political matters. At a time when the justiciability of constitutional conventions is being examined, now is a good moment to explore ouster clauses, being a point at which judicial involvement in political matters comes to a crux.

Commentary on ouster clauses is divided. Constitutionalists such as Professor Griffith identified ouster clauses as indispensable methods of protecting the sovereignty of Parliament from a mutinous judiciary. This perspective holds that ouster clauses are useful tools for ensuring that specialist bodies may be the final word in adjudication which is relevant to their expertise.[7] Differing interpretations of ouster clauses are demonstrative of the 'red-light'-'green-light' premise of Harlow and Rawlings on how policy factors affect application of legal principles. Harlow and Rawlings provide two ideological approaches to the rule of law: the 'red-light' approach is characterised by a laissez-faire distrust in the advancing power of the state; the 'green-light' approach sees state involvement in the life of individuals as a necessary prerequisite for meeting social goals.

For the former, ouster clauses prevent the legal right and duty of the courts to review the decisions of a state encroaching upon civil liberties; for the latter, ouster clauses are efficient methods of protecting social advancement from a conservative judiciary.[8] Griffith's concern is one that academics have recognised as belonging to the 'green-light' school: ouster clauses provide "consistency and finality" in legislative decisions.[9] Discussion on the law of ouster clauses generally

6 P Leyland and G Anthony, "Express and implied limits on judicial review: ouster and time limit clauses, the prerogative power, public interest immunity" in *Textbook on Administrative Law* (Oxford University Press 2016), 392.

7 JAG Griffith, *The Politics of the Judiciary* (Fontana Press 1997), 123–4, 340.

8 C Harlow and R Rawlings, *Law and Administration* (Cambridge University Press 1984).

9 *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 393.

orbits around two main conclusions. First, that those favouring the ouster clause do so from a 'green-light' perspective. Second, that the traditional position—which held clauses would suffice to oust the power of the courts to review an executive body if it acted within its jurisdiction—is now defunct. Through the previous case law and the Supreme Court judgement in *Privacy International*, this paper will examine these two claims to consider whether the judiciary's attitude towards ouster clauses has become an indurate legal principle.

It will be argued that, from the decision in *Privacy International*, these must be revised. First, that the new favourable interpretations of ouster clauses come from what might be considered an 'amber-light' approach, respecting liberal social mores whilst expressing concern with an encroaching executive. Second, that although the majority decision of the Supreme Court undermined the traditional distinction between ouster clauses concerning errors within jurisdiction and without jurisdiction, in practice this is still present. In qualifying its judgement in *Privacy International* as such, the Supreme Court decision has overturned an agreement struck between Parliament and the judiciary on sovereignty and jurisdiction. This change, it will be argued, is representative of the argument advanced by Poole in 2009 of the "reformation" in administrative law from the language of unreasonableness in *Associated Provincial Picture Houses Ltd v Wednesbury Corporation*[10] to rights-based judicial review. Where once generally ultra vires behaviour was the trigger for review of bodies behaving beyond their permitted jurisdictional bounds, now anthropomorphic rights-based arguments provide the foundation for judicial action.[11] It is in this conceptual matrix that the decision in *Privacy International* was decided against the traditional interpretation of the ouster clause. Although the judgement simplifies the law on ouster clauses, *Privacy International* has created the potential for greater confusion and constitutional ambiguity.

II. The concept of the ouster clause

It is useful to understand ouster clauses and judicial interpretations of them by cutting through their theoretical difficulties: what Beatson called an "intrinsic mosaic of conceptual formulae."[12] Ouster clauses are methods of precluding judicial review of determinations made by tribunals and lower courts. The logic behind them works thus: in a piece of legislation, Parliament bestows upon a Tribunal the jurisdiction to make decisions over hearings in a particular area.

[10] *Associated Provincial Picture Houses Ltd v Wednesbury Corporation* [1948] 1 KB 223.

[11] T Poole, "The Reformation of English Administrative Law" (2009) 68 CLJ 142, 142.

[12] J Beatson, "The Scope of Judicial Review for Error of Law" (1984) 4 OJLS 22, 31.

A section of the Act provides that the Tribunal has authority to determine in proceedings whether or not a third party meets statutory requirements.

Hypothetically, say Parliament deemed it appropriate to introduce a new oversight of the making and registration of cheese. The fictitious Dairy Products Act provides in section 1 that the Cheese Registration Tribunal has jurisdiction to determine whether or not a dairy product may be called a 'cheese'—depending on requirements such as ingredients and processes. The authority is given to the Tribunal because Parliament deems it to be an agency particularly qualified in the area of policy—the Investigatory Powers Tribunal, for example, consists of prominent judges and academics who are well-placed to make determinations on the security services. Due to this particular expertise, another section of the Act provides that any decision made by the Tribunal is not capable of judicial review. So, in the Dairy Products Act, section 2 states that 'the Cheese Registration Tribunal's determinations shall not be reviewable in any court': a total ouster clause.

Alternatively, section 2 of the theoretical Act might provide that legal action may be brought against a decision of the Tribunal, but only within six months of the determination being made. This is a time limit clause and, although the concept is slightly different to total ouster clauses, the courts generally consider the two together. Parliament relies on the Diceyan legal theory that it is the source of sovereignty which courts must respect.[13] Legislation must therefore be read in line with parliamentary intention. The courts, however, apply the fundamental principles of the rule of law in keeping governments in check. One aspect of this is that the courts have the right and duty to review an executive body if its actions exceed the legislative authority granted upon it by Parliament. Per Edin, as the courts interpret parliamentary intention, they will hold that Parliament would not have the intention to allow an executive body to violate legal principles without absolute clarity.[14] Therefore, if a Tribunal makes a decision which breaches the jurisdiction placed upon it by Parliament, the courts are considered competent to review it. To return to the cheese analogy: what would happen if the Tribunal decided that a coagulated dairy product, despite meeting the requirements and which is—for all intents and purposes—a cheese, should not be called a cheese? What about if, after this decision was made, it was revealed that members of the Tribunal had shares in a competitor's business? Or if it was held that the Tribunal's understanding of what could be considered a cheese was erroneous? Or if the applicants claimed that the Tribunal misapplied the statute to exclude them? Any

[13] AV Dicey, *An Introduction to the Study of the Law of the Constitution* (1885).

[14] DE Edin, "A Constitutional Right to Judicial Review: Access to Courts and Ouster Clauses in England and the United States" (2009) 57 AJCL 67, 68.

decision of the Tribunal that extended beyond its authority would undermine the effectiveness of the ouster clause in two ways.

First, the decision by the Tribunal may be erroneous as to undermine the words of the statute and remove legislative control, so that the courts may review it. Second, the decision by the Tribunal may be erroneous as to disrupt or threaten a key principle of English law, which the courts may review by right. The debate lies, however, in deciding what errors of law may amount to breaching the jurisdictional remit of a tribunal and therefore nullifying an ouster clause. In considering whether ousters are effective or not, the question is whether an error by the executive body invalidates its authority, and removes the statutory protection of the ouster clause.

III. The landmarks of the ouster clause

Before moving to the Supreme Court's judgement in *Privacy International*, it is prudent to first examine and understand the preceding case law. Much academic commentary has been written on these cases: the purpose of revisiting the judgements is not to question the decisions made, but to demonstrate the journey which has led to *Privacy International*. In 1980, after the decision in *Anisminic*, Professor Wade wrote that the judiciary "have almost given us a constitution, establishing a kind of entrenched provision to the effect that even Parliament cannot deprive them of their proper function."[15] Although this comment may seem overstated for the time, it is somewhat prescient of the decision in *Privacy International*, which has streamlined the jurisdiction of the courts to overturn ouster clauses in extending judicial authority to review errors. When reading the case law on ouster clauses, one must consider two distinct but related areas of discussion: first, the face-value legal jurisdiction of the court to review, and second, the political implications of the courts' review of executive decisions. Since the 1960s, there has been a slow progression towards judicial comfort in throwing out ouster clauses for any erroneous decision made by the executive body in question. Previously this would only stem from a legal error that went beyond the jurisdiction allotted to it by Parliament. With the authority of the Supreme Court judgement, and this distinction seemingly over, the road ahead may have fewer options for judges to take.

A. The old ouster clause

When ouster clauses were brought in front of the courts prior to the 1960s, there was a distinction drawn between two types of erroneous decisions that these

[15] HWR Wade, *Constitutional Fundamentals* (Stevens & Sons 1980) 68.

attempted to cover: non-jurisdictional (an executive body committing an error of law within the powers conferred upon it), and jurisdictional (an executive body committing an error of law outside the powers conferred upon it). Blackstone wrote that inferior courts were subject to supervision:

> [W]here they concern themselves with any matter not within their jurisdiction… or if in handling matters clearly within their cognizance they transgress the bounds prescribed to them by the laws of England… else the same question might be determined different ways, according to the court in which the suit is depending: an impropriety which no wise government can or ought to endure, and which is, therefore, a ground of prohibition.[16]

Although courts were capable of reviewing jurisdictional errors of law by virtue of the public bodies in question behaving ultra vires, an ouster clause would be effective in preventing judicial review of a non-jurisdictional error of law. The rationality for this is simple: if an executive body acts beyond its jurisdictional limits then the courts have an intrinsic right and duty to review that potentially unlawful action. This reasoning goes to the heart of the grounds for judicial review that would later become instilled in *Wednesbury* principles and the doctrine of ultra vires. In *R v Cheltenham Commissioners*,[17] the court prevented the success of an ouster clause in a jurisdictional case, in which Lord Denman CJ stated that: "The statute cannot affect our right and duty to see justice executed."[18] In this case, a paving and lighting Act for the town of Cheltenham enabled commissioners to raise money through a tax. Those taxed could appeal to the quarter sessions if they felt the rate affecting them was unreasonable, but the Act ousted the jurisdiction of the courts to review the sessions' decisions. Section 134 held that "the determination of the said justices in their said general quarter sessions, or adjournment thereof, shall be final, binding, and conclusive to all intents and purposes whatsoever."[19] Section 136 further enacted "that no order, verdict, rate, assessment, judgement, conviction, or other proceeding… shall be quashed or vacated for want of form only, or be removed or removable by certiorari, or any other writ or process whatsoever, into any of his Majesty's courts of record at Westminster…"[20] Any certiorari of the quarter sessions' decision on the rates was ousted by these provisions.

However, in an appeal against a rate in which the respondents objected to an admission of evidence, it was found that three of the eleven magistrates who held the evidence admissible had interests in the property concerned. It was held that the attempted ouster clause was undermined by the jurisdictional error of the

16 W Blackstone, *Commentaries on the Laws of England, book III* (1768), 112.

17 *R v Cheltenham Commissioners* [1841] All ER Reprints 301 [1835–1842] (QB).

18 *ibid*, 303 (Lord Denman CJ).

19 1 & 2 Geo 4, c cxxi, section 134.

20 *ibid*, section 136.

quarter session: the interests of the three magistrates in question meant that the proceedings were "improperly constituted" and that, as a result, sections 134 and 136 of the Act did not function. Lord Denman CJ summarised how the behaviour of the quarter sessions warranted the court's interference, stating that the clause could in no way "preclude our exercising a superintendence over the proceedings, *so far as to see that what is done shall be done in pursuance of the statute* [emphasis added]."[21]

The court's ability to undermine this ouster clause stemmed from the jurisdictional error that had been committed. The reasoning for which, however, was that of legislative supremacy as opposed to purely preventing the unlawful behaviour. Parliamentary and statutory authority still took precedence over judicial intervention: the logic of the court was to refuse the ouster clause so that the legislation itself was not undermined. Seeing justice executed was predicated on reviewing the quarter session's judgement, not because of the nature of that decision, but because of the way in which that decision was made. The ouster clause failed to prevent review in this jurisdictional error because the court was, per Williams J, "badly constituted".[22] The effectiveness of the clause was undermined by the quarter session's administrative failures.

Here, however, lie the limits of the old jurisdictional/non-jurisdictional divide: ouster clauses were rendered ineffective only by practical errors made by the executive body that meant they had behaved ultra vires. The courts would not be able to override an ouster clause if the error made did not take the body outside of its allotted power. In *Cheltenham Commissioners*, the corrupt behaviour of the three interested magistrates sitting in the quarter session may have made the body reviewable, but the actual determination reached by that local court alone would have had no justiciability.

For non-jurisdictional errors, therefore, the courts were not considered competent to override any ouster clauses or to review any executive bodies they covered. The case of *Smith v East Elloe Rural District Council*,[23] decided in a period of "judicial quietism",[24] saw the House of Lords interpret an ouster clause so broadly that not even corrupt action by the executive body was considered to have invalidated it. In the facts of this case, there was a challenge to a compulsory purchase order under the Acquisition of Land (Authorisation Procedure) Act 1946, which held that a challenge had to be made within six weeks of the order. If this deadline was not met, per paragraph 16 of schedule 1, the compulsory purchase order "shall not… be questioned in any legal proceedings whatsoever."[25]

[21] *Cheltenham Commissioners* (n 17) 303 (Lord Denman CJ).

[22] *ibid*, 305 (Williams J).

[23] *Smith v East Elloe Rural District Council* [1956] 1 All ER 855 (HL).

[24] *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 395.

[25] Acquisition of Land (Authorisation Procedure) Act 1946, schedule 1, part 4, para 16.

The order in question was made in 1948, against which Mrs Smith took action in 1954 on the grounds that it "had been made and confirmed wrongfully and in bad faith," and invalidated the clause to have no application in ousting the jurisdiction of the courts.[26]

This challenge against the district council failed. First, the appellant's lateness in bringing action was held to immediately breach the time limit clause, which was arguably fair. As Adler wrote, there were clear policy reasons for the time limit placed in the Act, and only a "conceptual purist" would be unhappy with this despite having an adequate remedy in place.[27] Second, in addressing the appellant's argument that the compulsory purchase order would only be valid if made in good faith, Viscount Simonds held it "impossible to qualify the words of the paragraph" as such. The view of his Lordship was that the pure reading of the words of the statute meant that any chance of undermining the ouster clause for fraud was impossible:

> It cannot be predicated of any order that it has been made in bad faith until it has been tested in legal proceedings, and it is just that test which paragraph 16 bars… What else can "compulsory purchase order" mean but an act apparently valid in the law, formally authorised, made, and confirmed?[28]

Under this authority a fraudulent act amounting to a jurisdictional error of law would be considered both valid and legal by the very nature of its being made. Any ouster clause could be inviolable, because the sovereignty of a Parliamentary direction would override any erroneous aspects of the determination made by an executive body: jurisdictional errors could be interpreted as non-jurisdictional errors and the courts' powers to review would be dead in the water. This was a judgement that favoured the orthodox view of parliamentary sovereignty and the purposive approach to the statutory intentions of the legislature. In doing so, it created a forceful rationale for ousting the courts' jurisdiction. Lord Morton took this further, writing that "it does not seem to me inconceivable, though it does seem surprising, that the legislature should have intended to make it impossible for anyone to question in any court the validity of a compulsory purchase order on the ground that it was made in bad faith."[29] Taking this judgement to its furthest theoretical conclusion, an executive body could do anything illegal or fraudulent,

[26] *Smith v East Elloe Rural* (n 23) 855D–E.

[27] J Adler, "Time Limit Clauses and Judicial Review – *Smith v East Elloe* Revisited" (1975) 38 MLR 274, 276.

[28] *Smith v East Elloe Rural* (n 23) 859A–B (Viscount Simonds).

[29] ibid 863C (Lord Morton).

and a sufficiently-drafted clause would succeed in precluding the court from its power to review. This would undermine a key principle of the rule of law as we understand it today.

The judgement in *Smith* was an extreme decision and not an authority that was followed for long, but its extremity is worth noting. It demonstrated that, in a post-war environment, the House of Lords was willing to protect executive decisions from the courts. This decision represents a conceptual definition that certain academics favoured in limiting the power of the courts to arbitrarily review. Taylor expressed disdain with the way that abuse of discretion was a treated like a "grab-bag" for some courts, "from which a ground of review can always be found to suit the conclusion sought to be reached on the merits."[30] At a time of social and industrial change, and representative of the 'green-light' interpretation, their Lordships were willing to limit individual protection from fraudulent executive behaviour in favour of government projects. It was a period in which the court held the concept of executive jurisdiction to be broad. Not only did the time limit clause have effect, but so too did the words of the Act to prevent review in a more general sense. In this intellectual environment, the onus was on the claimant—not the government—to prove that their reading of the clause had effect. In the dissenting judgement, Lord Reid held that the general words used in paragraph 16 did not exclude fraudulent action, but were "limited so as to accord with the principle" that "words are not to be read as enabling a deliberate wrongdoer to take advantage of his own dishonesty."[31] There was now a precarious relationship between the courts' jurisdiction to review and Parliament's capacity to oust justiciability, as the traditional distinction between legal errors was thrown out. It was in this uneasy judicial precedent that two formative cases on ouster clauses were decided.

B. Expanding justiciability: *Ex parte Gilmore and Anisminic Ltd*

The well-documented cases of *R v Medical Appeal Tribunal, ex parte Gilmore*[32] and *Anisminic* provided the leading authority on ouster clauses from the mid-twentieth century onward. *Ex parte Gilmore*, the leading judgement that holds finality clauses are not to be recognised by the courts as an effective method of ousting review, addressed legislation that sought to provide that no decision by the Medical Appeal Tribunal could not be challenged in a court.[33] In this case, section 36(3) the National Insurance (Industrial Injuries) Act 1946 stated that "any

[30] GDS Taylor, "Judicial Review of Improper Purposes and Irrelevant Considerations" (1976) 35 CLJ 272, 272.

[31] *Smith v East Elloe Rural* (n 23) 869D (Lord Reid).

[32] *R v Medical Appeal Tribunal, ex parte Gilmore* [1957] 1 QB 574 (CA).

[33] *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 393.

decision of a claim or question… shall be final",[34] covering the injury of a collier worker that left him blind. Although two of the three medical boards assessed the disablement at 100 per cent, one made no award, and on appeal the Tribunal assessed the damage as 20 per cent. Mr Gilmore applied for a certiorari against this decision by claiming that it failed to meet the requirements of regulation 2(5) of the National Insurance (Industrial Injuries) (Benefit) Regulations 1948.[35]

Surely finding for this appeal and allowing the certiorari would fly in the face of direction of their Lordships in *Smith*? At its face, this could be conceived as a non-jurisdictional error, as the tribunal was within its powers to reach the conclusion that it did. Even when interpreted as a jurisdictional error, the House of Lords had held that clear wording such as section 36(3) would make a tribunal decision unreviewable by the courts. Denning LJ characteristically turned to authorities over three hundred years old to hold that the clause was ineffective to oust the courts' ability to issue a certiorari and revisited the decision of the Medical Appeal Tribunal: "I find it very well settled that the remedy by certiorari is never to be taken away by any statute except by the most clear and explicit words."[36]

According to Denning LJ, section 36(3) did not suffice to be 'clear and explicit', and the court called the House of Lords' bluff. The judgement found for the traditional distinction between jurisdictional and non-jurisdictional errors, identified the facts of this as jurisdictional, and made it clear that the words of the statute were not sufficient to oust the right of certiorari. Denning LJ held the decision to be a jurisdictional error of law due to the incorrect application of a medical report by the Tribunal, so that it was "open for this court to issue a certiorari to quash it for an error of law on the face of the record."[37] Therefore, as a traditionally jurisdictional error of law, the statute would not be enough to oust the review of the court: "The word "final" is not enough. That only means "without appeal." It does not mean "without recourse to certiorari." It makes the decision final on the facts, but not final on the law."[38]

Just months after the disruption to the traditional distinction in the judgement of *Smith*, the Court of Appeal repaired and reused it. In an effective history lesson, Denning LJ then explained why legislation that sought to quash the right of certiorari should not be effective in modern hearings. In earlier centuries certiorari was used too liberally by the courts "to quash for technical defects" instead of defeating "a substantial miscarriage of justice."[39] Now that

[34] National Insurance (Industrial Injuries) Act 1946, s 36(3).
[35] *ex parte Gilmore* (n 32) 575.
[36] *ex parte Gilmore* (n 32) 583 (Denning LJ).
[37] *ibid* 585 (Denning LJ).
[38] *ibid* 583 (Denning LJ).
[39] *ibid* 586 (Denning LJ).

this was no longer the case, the need to protect the executive from unnecessary orders of certiorari was gone, and clauses could not be effective in ousting the "ancient writ".[40] The decision in *ex parte Gilmore* was representative of the policy considerations that lie behind ouster clauses just as much as the legal reasoning. Per Griffiths, later judgements would come to rely on this subversion of the rule in *Smith*, by reaffirming the position on finality clauses and jurisdictional errors.[41] The judgement was important in maintaining the traditional distinction between errors of law and, moreover, providing the factual and legal logic for the restrictions on ouster clauses. It was not until a decade later, however, that the House of Lords would once again comment on the issue.

The landmark case of *Anisminic* turned on a loss of property from the British mining company during the Suez Crisis in 1956, which was sold on to the Egyptian company TEDO. Anisminic Ltd were deeply dissatisfied with the compensation of £500,000 in addition to the property being sold at less than real value. In 1959, a treaty was agreed between the UK and Egypt to provide £27.5 million for any British property confiscated, for which distribution responsibility was vested in the Foreign Compensation Commission.[42] The Commission operated under the Foreign Compensation (Egypt) (Determination and Registration of Claims) Order 1962 which held a claim could be established if, (a) the applicant was the person referred to as owner of the property, and (b) the person referred to *and* any person who became successor in title were British nationals.[43] This Order, in turn, was governed by the Foreign Compensation Act 1950 section 4(4) which provided: "The determination by the commission of any application made to them under this Act shall not be called in question in any court of law."[44] When the Commission held that Anisminic Ltd failed in its application for compensation because TEDO was not a British national, the company argued that the Commission had misunderstood the Order, and the House of Lords had to consider whether section 4(4) of the Foreign Compensation Act did oust the courts' power to review.[45] Following the reasoning in *Smith*, it would do so because the process of discovering whether the Commission had acted ultra vires and misunderstood the Order would be excluded by the statute.

However, in the majority their Lordships rejected this line of argument, with Lord Reid stating that: "It is a well-established principle that a provision

40 *ibid*.

41 J Griffiths, "Judicial Review for Jurisdictional Error" (1978) 38 CLJ 11, 11–2.

42 *Anisminic* (n 2) 164D–F.

43 Foreign Compensation (Egypt) (Determination and Registration of Claims) Order 1962 (SI 1962 No 2187).

44 Foreign Compensation Act 1950, section 4(4).

45 *Anisminic* (n 2) 164F–G.

ousting the ordinary jurisdiction of the court must be construed strictly—meaning, I think, that, if such a provision is reasonably capable of having two meanings, that meaning shall be taken which preserves the ordinary jurisdiction of the court."[46]

Righting the perceived wrong in *Smith* that Lord Reid was a dissenting voice against thirteen years earlier, the House of Lords expanded the definition of what a jurisdictional error might consist of. It was held that "there are many cases where, although the tribunal had jurisdiction to enter on the inquiry, it has done or failed to do something in the course of the inquiry which is of such a nature that its decision is nullity."[47] In this reasoning, therefore, a tribunal or other executive body could act well within its allotted jurisdiction but do something that would compromise the decision. Giving a non-exhaustive list of what this something might include; his Lordship took great pains to insist that this did not extend the definition of jurisdictional error into the non-jurisdictional realm.[48]

However, not everyone has found this convincing: Griffiths wrote that "the decision effectively obliterated the distinction between errors of law made within, as distinct from in excess of, jurisdiction."[49] Reminiscent of *Cheltenham Commissioners*, and defining a typically non-jurisdictional error as a jurisdictional one due to marred decision-making, the decision flipped the reasoning in *Smith* on its head. Where Viscount Simonds had expanded non-jurisdictional error, Lord Reid used a mirrored reasoning to expand jurisdictional error, into which fell the purported misunderstanding by the Commission of the Order's nationality requirement. Dissenting, Lord Morris held this unnecessary: the Act, he stated, did not prevent inquiries "to decide whether the commission has acted within its authority or jurisdiction."[50] His Lordship expressed concern that the majority judgement would provide the impetus for judicial intervention in previously protected situations.[51]

As Leyland and Anthony have highlighted, this reasoning protects the basic right of the court to review from total exclusion, as the judge sitting on the application for judicial review will determine whether the decision is valid.[52] This must be considered when examining the decision in *Privacy International* because, per Griffiths, although the House of Lords "took a very generous view of the jurisdictional error in the circumstances of that case, each of the majority judgements in that decision insisted that an error of law could still occur within

46 *ibid* 169B (Lord Reid).
47 *ibid* 170B–D (Lord Reid).
48 *ibid*.
49 Griffiths, "Judicial Review for Jurisdictional Error" (n 41) 11.
50 *Anisminic* (n 2) 179D (Lord Morris).
51 *ibid* 180D–G (Lord Morris).
52 *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 396.

jurisdiction."[53] The language adapted, so that decisions outside of the body's jurisdiction became 'purported' determinations, and any legal error would be capable of review. Just two years earlier, Taylor had advocated for further entrenchment of the distinction between grounds, which he argued would clarify the law on judicial review of improper purposes.[54] *Anisminic*, however, decided that clarity was to be found in a conceptual opposite. The decision of the House of Lords expanded the errors capable of disabling an ouster clause, but did so under the one title of jurisdiction. In the decades following, the courts applied the reasoning of *Anisminic* in a consensus that, whilst legitimising the capability of certain ouster clauses, shifted the realm of jurisdictional error.

C. The post-*Anisminic* consensus

In the decades leading up to *Privacy International*, judgments followed a consensus laid down in *Anisminic*. Although some cases on ouster clauses reached the House of Lords, and later the Supreme Court, none really called into question the ambit of jurisdictional error. In *Pearlman*, the Court of Appeal considered whether a judge's decision was open to certiorari, concerning the Housing Act 1974. Schedule 8 to the Act held that "works amounting to structural alteration, extension or addition" to a rented house by a tenant entitled them to a reduction in rates.[55] The judge held that the tenant's own installation of modern central heating to the flat did not amount to such works under Schedule 8.[56] Moreover, concerning the decision of the judge, paragraph 2(2) of the Schedule provided that "any such determination shall be final and conclusive."[57] The court was also asked to consider the authority of the County Courts Act 1959, which provided that "no judgement or order of any judge or county courts… shall be removed by appeal, motion, certiorari or otherwise into any other court whatsoever…"[58]

Lord Denning MR, Lane and Eveleigh LJJ all held that the judge erred in his decision that the works did not amount to structural alteration; Lord Denning MR and Eveleigh LJ went further, and applied the decision in *Anisminic* to state that this error was one "in law", so that the judge "wrongly deprived himself of jurisdiction" and paragraph 2(2) of the Schedule failed to exclude certiorari.[59] As the case concerned an error made on a point of law, Lord Denning MR held

[53] Griffiths, "Judicial Review for Jurisdictional Error" (n 41) 14.
[54] Taylor, "Judicial Review of Improper Purposes and Irrelevant Considerations" (n 30) 291.
[55] Housing Act 1974, schedule 8, para 1(1).
[56] *Pearlman* (n 4) 736H–7C.
[57] Housing Act 1974, schedule 8, para 2(2).
[58] County Courts Act 1959, section 107.
[59] *Pearlman* (n 4) 737C–F.

that section 107 of the County Courts Act did not remove the right of certiorari, because the judge had acted without jurisdiction. He then went further, to suggest that the traditional jurisdictional/non-jurisdictional distinction defunct, breaching the decision in *Anisminic* by removing any erroneous determination in which an ouster clause may have effect.[60]

This was not a popular decision, which revealed the fragility of the consensus created by *Anisminic*: as Griffiths wrote, it "threaten[ed] to expose the courts to a direct confrontation with Parliament."[61] Even if the distinction between jurisdictional and non-jurisdictional errors existed in name only, it was held to be a principle central to the rule of law and separation of powers, that underpin the relationship between the courts and the Parliament. Parliamentary intention to allow a body to err within its own jurisdiction is still a valid intention; for the courts to ignore that would undermine the very notion of parliamentary sovereignty. On this matter, Lane LJ dissented strongly from the majority, stating that "if this judge is acting outside his jurisdiction, so then is every judge who comes to a wrong decision on a point of law."[62]

The majority judgement in *Pearlman* extended the judicial power beyond what *Anisminic* had permitted and, soon after in *Re Racal Communications*,[63] the House of Lords returned to the consensus. Their Lordships held that section 441(3) of the Companies Act 1948, which stated that a decision by a High Court judge "shall not be appealable",[64] was effective under *Anisminic*. The decision was two-fold: first, that in considering the case the Court of Appeal had acted beyond its own jurisdiction, and was therefore itself erroneous; second, that although errors of law are reviewable, where a power is granted to a tribunal it is not for the courts to assume its limits: "there was no presumption that Parliament did not intend to confer on it a power to determine questions of law going to its jurisdiction as well as questions of fact".[65] What the majority decision in *Pearlman* had read as an automatic authority of the courts to review anything erroneous was now limited by the House of Lords. The consensus, therefore, was one of complex logic that found executive bodies to be justiciable on a case-by-case basis. Per Lord Diplock,

[60] *ibid*, 743G–4G (Lord Denning MR).

[61] Griffiths, "Judicial Review for Jurisdictional Error" (n 41) 14.

[62] *Pearlman* (n 4) 750A (Lane LJ).

[63] *Re Racal Communications* [1981] AC 374 (HL).

[64] Companies Act 1948, section 441(3).

[65] *Re Racal Communications* (n 63) 375E–G.

when legislation "provides that the judge's decision shall not be appealable, they cannot be corrected at all."[66]

Following the authority of *Re Racal Communications*, the distinction still existed, but on the precarious understanding that executive bodies may be capable of determining their own reviewability if the statute confers such purview upon them. Following this, in *R v Lord President of the Privy Council, ex parte Page*,[67] the nature of a legal error was considered with a more conservative understanding of what powers fell under jurisdiction. The House of Lords held that the Lord President of the Privy Council – acting as visitor on behalf of Hull University – was within their powers to reject a petition for review. The appellant argued that, because he had been laid off without "good cause" per the requirements of the university statutes, this was a jurisdictional error under *Anisminic*. Lord Browne-Wilkinson simply stated that this was an error of law which was covered within the discretion of the visitor; as such, it was held to be a traditionally non-jurisdictional case, and the court's ability to review was limited. He noted that the decision in *Anisminic*:

> [R]endered obsolete the distinction between errors of law on the face of the record and other errors of law by extending the doctrine of ultra vires… Parliament had only conferred the decision-making power on the basis that it was to be exercised on the correct legal basis: a misdirection in law in making the decision therefore rendered the decision ultra vires.[68]

The judgement in *ex parte Page* ascertained that a jurisdictional error did not occur automatically unless an executive action question had been ultra vires. This in turn expanded the scope of what might be considered a body's jurisdiction. In the Court of Appeal case of *R v Secretary of State for the Home Department, ex parte Fayed*,[69] two brothers who applied for British naturalisation were refused, a decision for which the Home Secretary refused to give any reasons. Under the British Nationality Act 1981, the Home Secretary was not "required to assign any reason for the grant or refusal of any application," which would "not be subject to appeal to, or review in, any court".[70]

Lord Woolf MR held, however, because no explanation was given, the brothers were not given the legal fairness to which they were entitled.[71] It was a

[66] *ibid* 384F (Lord Diplock).
[67] *R v Lord President of the Privy Council, ex parte Page* [1993] AC 682 (HL).
[68] *ibid* 701F (Lord Browne-Wilkinson).
[69] *R v Secretary of State for the Home Department, ex parte Fayed* [1997] 1 All ER 228 (CA).
[70] British Nationality Act 1981, section 44(2).
[71] *ex parte Fayed* (n 69) 242B (Lord Woolf MR).

purely jurisdictional error that had been made, and as a result, the decision was open to review. This decision denotes a period in which, as Poole has identified, the Human Rights Act 1998 began to take precedence and introduced "a structure in which the responsibilities that public authorities have in relation to rights" for the courts to reinforce.[72] Slowly, an individual rights-based concept of judicial review replaced the purely ultra vires position advanced in *Anisminic*.

In *R (on the application of Cart) v Upper Tribunal*[73] the Supreme Court upheld the decision in *Anisminic*, citing the rule of law for holding judicial review of the Upper Tribunal available when the challenge raises a point of principle or if there is a compelling reason to hear the claim.[74] In this case, the Supreme Court was required to consider the Tribunals, Courts and Enforcement Act 2007, which provided that there was a right to appeal "any point of law arising from a decision made by the Upper Tribunal other than an excluded decision".[75] Did this suffice to oust judicial review of unappealable, 'excluded' decisions in the Upper Tribunal?

Lady Hale held that it did not: first, because there was a lack of the clear words required to make an ouster clause effective; second, that it was illogical for the statute to distinguish judicial review in certain situations which were "gathered together" in the legislation; third, that the rule of law requires tribunals and courts, which "are Parliament's bidding", to be justiciable.[76] Applying the decision provided in *Anisminic*, the judgement held that the Upper Tribunal had to be reviewable, in order to ensure that bad law did not become entrenched in the then-new tribunal system. Lord Phillips stated that reviewability was required "to guard against the risk that errors of law of real significance slip through the system."[77] The judgment in *Cart* conceptualised errors of law in a more modern framework. This was a broad definition, which made decision or determination with the potential to affect future cases capable of judicial review.[78]

However, this decision still turned on how legal errors fell outside the jurisdiction of the system. As the Upper Tribunal is an appellate court, questions of legal points and principle are reviewable, because an error of law would take the body out of its jurisdiction. The distinction that was applied in the earlier case

72 Poole, "The Reformation of English Administrative Law" (n 11) 166.

73 *R (on the application of Cart) v Upper Tribunal* [2011] UKSC 28.

74 *ibid* [37] (Lady Hale).

75 Tribunals, Courts and Enforcement Act 2007, section 13(1)

76 *Cart* (n 73) [37] (Lady Hale).

77 *ibid* [91] (Lord Phillips).

78 E Craven, "Case Comment: *R (Cart) v The Upper Tribunal; R (MR (Pakistan)) v The Upper Tribunal (IAC)* [2011] UKSC 28" (*UKSC Blog*, 4 July 2011) <http://ukscblog.com/case-comment-r-cart-v-the-upper-tribunal-r-mr-pakistan-fc-v-the-upper-tribunal-immigration-asylum-chamber-2011-uksc-28/> (accessed 20 July 2019).

law exists in this judgement, with the Supreme Court emphasising that the danger of erroneous law becoming stuck in the decisions of the Upper Tribunal prevents an ouster clause from being effective, as it goes to the heart of the purpose of the Upper Tribunal. The consensus established in *Anisminic* is one that respected the sovereignty of Parliament whilst ensuring that individuals had the right to judicial protection against an overbearing executive. The danger with increasing the definition of jurisdictional error, as seen in the decision in *Pearlman*, is that it allowed the courts to conflate law and fact to create situations in which ouster clauses would be ineffective. As Beatson has noted, such an administrative legal system does not use the concept of jurisdictional error, but makes error of law "a façade" for a system that favours potentially arbitrary decisions as to whether the courts may review.[79] What followed *Anisminic* was a series of attempts by the courts to ensure that all legal errors were justiciable whilst conforming to Diceyan constitutional theory on Parliamentary sovereignty. The judgement in *Privacy International* must be read as part of this legal history.

IV. The *Privacy International* case

As this paper has demonstrated, prior to the decision in *Privacy International* there was a period of general stability in which it was held that in order to be considered jurisdictional, an error must be one that is fatal to the very purpose of the executive body in question. The question in *Privacy International* turned on whether an error of determination, subject to the influence of *Anisminic* and a new conceptual environment that emphasised rights-based review.

A. The facts

The Investigatory Powers Tribunal (IPT) was established to hear complaints relating to the use of investigatory powers by the intelligence services. It was set up under the Regulation of Investigatory Powers Act 2000, section 67(8) which provided that: "Except to such extent as the Secretary of State may by order otherwise provide, determinations, awards, orders and other decisions of the Tribunal (including decisions as to whether they have jurisdiction) shall not be subject to appeal or be liable to be questioned in any court."[80]

The IPT allowed the hacking of computers under the Intelligence Services Act 1994, which provides that the Secretary of State may issue a warrant "authorising the taking of such action as is specified in the warrant in respect of any property so

[79] Beatson, "The Scope of Judicial Review for Error of Law" (n 12) 43.

[80] Regulation of Investigatory Powers Act 2000, section 67(8).

specified".[81] Advocating that they were the victims of unlawful surveillance by the intelligence services, Privacy International argued this behaviour was incompatible with Convention rights and amounted to an error of law. This challenge failed before the IPT, which held that section 5 of the Intelligence Services Act allowed the surveillance to take place, and Privacy International sought judicial review.[82] Two issues went before the court: first, whether the Act was effective in ousting the jurisdiction of the court to review the Investigatory Powers Tribunal for a legal error, and second, whether Parliament can oust the jurisdiction of the court to quash the decision of an inferior court or tribunal.[83] Underlying this was the issue of whether section 67(8) of the Regulation of Investigatory Powers Act was of equivalent effect to section 4(4) of the Foreign Compensation Act, which the House of Lords had deemed insufficient in *Anisminic* to oust judicial review.

B. The judgements of the divisional court and the court of appeal

Both the Divisional Court and the Court of Appeal held that section 67(8) of the Regulation of Investigatory Powers Act was effective in precluding judicial review of the IPT. In 2017, the Divisional Court stated that because the tribunal was already in a position of exercising supervision over the intelligence services, the facts of the case were in contrast with *Anisminic*. The role of the tribunal, Sir Brian Leveson P held, was that of expert oversight in quasi-judicial form. As a result, judicial review of the decision was effectively ousted by the Act, because the decision of the IPT had not breached its jurisdiction. He held that:

> There is a material difference between a tribunal—such as the Foreign Compensation Commission whose "determination" was in issue in *Anisminic*, SIAC, or the Upper Tribunal (when dealing with appeals from the First-tier Tribunal)—which is adjudicating on claims brought to enforce individual rights and the IPT which is exercising a supervisory jurisdiction over the actions of public

[81] Intelligence Services Act 1994, section 5.

[82] A Tucker, "Parliamentary Intention, Anisminic, and the Privacy International Case (Part One)" (*UK Constitutional Law Blog*, 18 December 2018) <https://ukconstitutionallaw.org/2018/12/18/adam-tucker-parliamentary-intention-anisminic-and-the-privacy-international-case-part-one/> (accessed 20 July 2019).

[83] *Privacy International* (SC) (n 1) [21] (Lord Carnwath).

authorities.[84]

As Daly noted, this line of argument made review of the IPT a logical difficulty: to review the IPT would be akin to judicial review of judicial review.[85] Leggatt J, however, held that the relevance of *Anisminic* lay with the traditional distinction of jurisdiction and protected the IPT from review in this case. *Anisminic* "decided that any determination based on an error of law, whether going to the jurisdiction of the tribunal or not, was not a "determination" within the meaning of the statutory provision."[86] Therefore, the decision of the IPT was not one that lay beyond the scope of its jurisdictional power, and as such the ouster clause was effective. For both judges the IPT was protected from review, but for considerably different reasons.

In the Court of Appeal, Sales LJ held that the language and context of the case differed from *Anisminic*. The Court held that, in the Act, "the word "decision" is stated to include a decision which (if judicial review or an appeal were available) might be found to have been made without jurisdiction because of an error of law on the part of the IPT".[87] According to Sales LJ, the purported determination that *Anisminic* turned on did not exist in this case, because section 67(8) of the Act had sufficiently extended the jurisdiction of the Tribunal. This prevented the Court of Appeal from doing what the House of Lords had done in *Anisminic*—finding purported decisions to lie outside the discretion of the Tribunal—by bringing them into its jurisdiction. The Court of Appeal held that this was effective due to the "very high quality" of the IPT, designed by Parliament to have expertise and independence in the delicate matter of national intelligence.[88]

This decision turned on the statutory differences between the Foreign Compensation Act and the Regulation of Investigatory Powers Act, and the role of Parliamentary intention in creating the IPT's authority over controversial security issues. Policy played just as big a role in the decision of the Court as the legal issues

[84] *R (on the application of Privacy International) v Investigatory Powers Tribunal* [2017] EWHC 114 (Admin) (DC) [42] (Sir Brian Leveson P).

[85] P Daly, "Ousting the Jurisdiction of the Courts: R (*Privacy International) v Investigatory Powers Tribunal* [2017] EWHC 114 (Admin)" (*Administrative Law Matters*, 6 February 2017) <https://www.administrativelawmatters.com/blog/2017/02/06/ousting-the-jurisdiction-of-the-courts-r-privacy-international-v-investigatory-powers-tribunal-2017-ewhc-114-admin/> (accessed 20 July 2019).

[86] *Privacy International* (DC) (n 84) [55] (Leggatt J).

[87] *R (on the application of Privacy International) v Investigatory Powers Tribunal* [2017] EWCA Civ 1868 (CA) [34] (Sales LJ).

[88] *ibid* [38], [42] (Sales LJ).

did. As Elliot has written, such scope means that the actual issue of the ouster clause in this case is "merely the tip of the constitutional iceberg".[89]

C. The judgement of the supreme court

Just months previously the Supreme Court had been asked to consider, in the case of *Lee v Ashers Baking Company Ltd and others*,[90] whether a decision of the Court of Appeal of Northern Ireland was precluded from appeal by devolution. In that judgement, Lord Mance distinguished between decisions of a judicial body and an administrative tribunal, and held that an ouster clause would be sufficient in principle to exclude an appeal on merits but not on procedural error.[91] In *Privacy International*, the Supreme Court took a different view. The judgement began from the point that the jurisdictional/non-jurisdictional divide is irrelevant: now, Lord Carnwath held, the starting assumption is that all errors of law are subject to review in a nuanced approach.[92] As this paper has shown, however, the hangover of the distinction between jurisdictional and non-jurisdictional errors is still felt, as ouster clauses attempt to expand a tribunal's jurisdiction over 'purported' decisions. The majority held that an ordinary reading of section 67(8) in the language of *Anisminic* failed to oust reviewability because "a decision which is vitiated by error of law, whether "as to jurisdiction" or otherwise, is no decision at all."[93] Tucker has argued that Parliament is aware of this: the authority of *Anisminic*, he argued, means that to be successful in drafting an efficient ouster would involve writing a clause that framed its intentions to protect it from a hostile interpretation by the courts.[94]

In *Privacy International*, therefore, the clause failed to oust the power of the courts to review because legal errors would fundamentally undermine it. Although the IPT is a specialist tribunal, Lord Carnwath held that this does not exempt it from justiciability, because the rule of law requires that it "conforms to the general law of the land."[95] It is important not to underestimate this decision: neither statutory expression, nor policy concerns, can protect a tribunal from review. Moreover, considering the second issue, the majority held that: "In all cases, regardless of

[89] M Elliott, "*Privacy International* in the Court of Appeal: *Anisminic* distinguished – again" (*Public Law for Everyone*, 26 November 2017) <https://publiclawforeveryone.com/2017/02/10/distinguishing-anisminic-ouster-clauses-parliamentary-sovereignty-and-the-privacy-international-case/> (accessed 20 July 2019).

[90] *Lee v Ashers Baking Company Ltd and others* [2018] UKSC 49.

[91] *ibid* [86], [88] (Lord Mance).

[92] *Privacy International* (SC) (n 1) [38] (Lord Carnwath).

[93] *ibid* [109] (Lord Carnwath).

[94] Tucker, "Parliamentary Intention, Anisminic, and the Privacy International Case (Part One)" (n 82).

[95] *Privacy International* (SC) (n 1) [139] (Lord Carnwath).

the words used, it should remain ultimately a matter for the court to determine the extent to which such a clause should be upheld".[96] By extending the decision in *Anisminic* to cover the IPT, hopes for ouster clauses are limited without expert statutory manipulation.

In response, Lord Sumption's dissenting judgement recognised the presence of the distinction between "errors of law going to jurisdiction and errors of law within jurisdiction", but held that the preceding case law had removed such subtleties from the judgement in *Anisminic*.[97] Although access to a court to review the legality of executive acts is essential to the rule of law, Lord Sumption held that "a right of appeal from such a body or right to call for a review of its decisions" is not.[98] If a body is by its nature expertly judicial, upon which Parliament has given authority, then decisions covered within its field ought to be conclusive. Lord Sumption provided the reasoning that, given the direction in *Anisminic* which had considered any error to exceed jurisdiction, Parliament had worded section 67(8) to cover the "merits" of the IPT's decisions.[99] His Lordship held that the Act had statutory clarity so to remove justiciability of the IPT's determinations, but not any procedural errors.[100] Furthermore, Lord Wilson held the legal authority of the IPT akin to the High Court in *Re Racal Communications*, the county court in *Pearlman* and the Upper Tribunal in *R (on the application of Cart)*, all of which made errors of law in unappealable decisions.[101] The dissenting judgements found in favour of the traditional jurisdictional distinction, and held the Tribunal to be protected by the ouster clause.

V. The ouster clause, ousted?

Reading the decision in *Privacy International*, it is possible to interpret this broad application of the decision in *Anisminic* as preventing any future ouster clause from being effective. The expansion of jurisdictional error now covers in essence every and any determination by a tribunal with which an appellant is dissatisfied, by arguing that there is an error of law, and it will then be for the court to decide upon. A common theme in judgements on ouster clauses is a complaint that Parliament has not been sufficiently clear in its language to oust the jurisdiction of the courts, as is the case in *Privacy International*.[102] However, it is difficult to see how

[96] *ibid* [144] (Lord Carnwath).
[97] *ibid* [181] (Lord Sumption).
[98] *ibid* [182] (Lord Sumption).
[99] *ibid* [201] (Lord Sumption).
[100] *ibid* [205] (Lord Sumption).
[101] *ibid* [244]–[252] (Lord Wilson).
[102] *ibid* [111] (Lord Carnwath).

Parliamentary intention could be any clearer than section 67(8) of the Regulation of Investigatory Powers Act 2000. The legal protection that would be afforded to the IPT due to its expert opinion and judicial authority, too, has been completely undermined: if the courts lack confidence in the IPT's jurisdiction over its expertise in intelligence services, then there is little hope for any other specialised tribunal. As Elliott noted, "the factual matrix presented in *Privacy International* is some way from the most egregious forms of ouster",[103] and yet it has been judicially cauterised.

There are two things to take from this new direction. First, that this is perhaps an unsurprising judgement, being the inevitable result of an expansion of jurisdictional error since the decision of the House of Lords in *Anisminic*. With the attempts to end the distinction between jurisdictional and non-jurisdictional error, as Leyland and Anthony noted and this paper has plotted, the courts held that any error of law is reviewable.[104] Surely to no sane person would this be seen as a negative development in the rule of law; indeed, it was part and parcel of the development of a clear principle of ultra vires. As Forsyth wrote, without ultra vires, "ouster clauses would be much more effective at excluding judicial review than they are today."[105] However, the decision in *Privacy International* may go too far, so that not only is any error of law reviewable, but so is any perceived error of fact. In removing the traditional distinction between jurisdictional and non-jurisdictional error, any error is reviewable, and the protection afforded to executive bodies in English constitutional conventions is undermined. What is the point, one might ask, of an expert tribunal, if their expertise is considered erroneous by a court and can be called into question? Although there is no legal principle to prevent Parliament drawing up an ouster clause, logically, their effectiveness is now limited by the conflation of errors within jurisdiction and errors without.

Second, that despite clarifying the law on ouster clauses, the Supreme Court judgement has the potential to cause conflict with Parliamentary authority by such expansion. Academics such as Tucker have held that the decision in *Anisminic* is emblematic of a common law doctrine of interpretation in which the court may depart from parliamentary intention when assessing ouster clauses. This is not new, but extends from a principle of "interpretive hostility" of intention in decisions

[103] Elliott, "*Privacy International* in the Court of Appeal" (n 89).

[104] *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 393.

[105] C Forsyth, "Of Fig Leaves and Fairy Tales: The Ultra Vires Doctrine, the Sovereignty of Parliament and Judicial Review" (1996) 55 CLJ 122, 128.

such as *R v Secretary of State for Transport, ex parte Factortame*,[106] and the requirements of the Human Rights Act 1998.[107]

The decision in *Privacy International*, therefore, is merely indicative of a judicial right and duty to ensure that the executive does not behave ultra vires. This is a sound argument until one considers the facts of the *Privacy International* case and the jurisdiction of the IPT: Parliament had vested a decision-making power concerning the intelligence services on the Tribunal. An "interpretive hostility" towards parliamentary intention in this case meant that the Supreme Court made a policy decision by holding a determination of the Intelligence Services Act by the IPT to be erroneous and disabling the ouster clause. Not only that, but the Supreme Court made a policy decision in a delicate political matter of national security. Winterton explained why, with ouster clauses, the buck must always stop with parliamentary sovereignty:

> Because ours is a government under law and the courts are the guardians of the law, those who seek to limit or oust judicial review should bear the burden of proving that in the particular case it is necessary or advisable to do so, and that an effective alternative form of review has been provided, but in a democracy the decision must be made by *Parliament*, not by the courts.[108]

Parliament, for better or worse, decided that the expertise of the IPT sufficed to prove that the ouster clause in section 67(8) was necessary in this case. As Lord Sumption held, the IPT was provided with "prescribed area of competence", which the facts of the case did not breach.[109] This is unsurprising, given his Lordship's numerous comments on the delicate relationship between Parliament and the judiciary, and his concerns with judicial resolution to political issues.[110] The elephant in the room of the judgement is the purpose of the Regulation of Investigatory Powers Act, which covers communication interception, use of data, and intrusive surveillance. These issues carry the sort of political complexity in which, for practical reasons, the courts previously would respect the authority of

[106] *R v Secretary of State for Transport, ex parte Factortame* [2000] 1 AC 524 (HL).

[107] Tucker, "Parliamentary Intention, Anisminic, and the Privacy International Case (Part One)" (n 82).

[108] G Winterton, "Parliamentary Sovereignty and the Judiciary" (1981) 97 LQR 265, 266.

[109] *Privacy International* (SC) (n 1) [211] (Lord Sumption).

[110] Lord J Sumption, "The Limits of Law" 27 Sultan Azlan Shah Lecture (20 November 2013).

ouster clauses.[111] In *A v Secretary of State for the Home Department*[112] Lord Bingham drew the line between political and judicial issues:

> The more purely political (in a broad or narrow sense) a question is, the more appropriate it will be for political resolution and the less likely it is to be an appropriate matter for judicial decision. The smaller, therefore, will be the potential role of the court. It is the function of political and not judicial bodies to resolve political questions.[113]

The Supreme Court judgement in *Privacy International*, however, presents itself as a modern day *Entick v Carrington*[114] protecting English liberties from an invasive political class and its computerised trespasses. The traditional 'green-light'-'red-light' interpretation of pro- and anti-ouster clause ideological camps is somewhat defunct in this judgement. Those opposed to the ouster clauses in *Privacy International* do so from a middle ground, as the consistency of the Tribunal's expertise in this controversial area of law does not suffice to oust judicial review of an intrusive executive: an 'amber-light' perspective.

This new perspective is one legitimised by the gradual move of judicial review from an administrative authority to a protector of rights. As Poole identified, the Human Rights Act 1998 has allowed a "new order" of judicial review to become the norm, the features of which are a "lack of any built-in limit to the proportionality test" in order to protect the rights of claimants.[115] In the decision of the Supreme Court in *Privacy International*, it is evident that this rights-based definition of judicial review has expanded, so that it now covers judicial interpretations of executive jurisdiction. In this conceptual environment an error of law undermines an ouster clause from precluding review, not because of administrative mistake, but because it undermines individual rights. The danger of this is a potential to cross into the political sphere. Where the ultra vires doctrine provided as clear boundaries to justiciable behaviour, a rights-based view of the ouster clause allows the courts to intrude on political decisions. In *Gibson v Lord Advocate*[116] Lord Keith's assessment of the separation of powers demonstrated why this might be the case. He held that:

> The making of decisions upon what must essentially be a political matter is no part of the function of the Court, and it is highly

[111] *Textbook on Administrative Law*, "Express and Implied Limits on Judicial Review" (n 6) 402.

[112] *A v Secretary of State for the Home Department* [2005] 1 AC 68 (HL).

[113] *ibid* [29] (Lord Bingham).

[114] *Entick v Carrington* (1765) 2 Wils KB 275.

[115] Poole, "The Reformation of English Administrative Law" (n 11) 145–6.

[116] *Gibson v Lord Advocate* [1975] SC 136 (CSOH).

> undesirable that it should be. The function of the court is to adjudicate upon the particular rights and obligations of individual persons, natural or corporate, in relation to other persons or, in certain circumstances, to the State.[117]

The decision in *Privacy International* expands the jurisdiction of the court itself into policy issues under the title of rights-based review. In so doing, decisions of a political matter have been made by the judiciary. The boundaries of executive jurisdiction in relation to the ouster clause have been redefined by the Supreme Court.

It is not improbable that the reader may feel this is a rather alarmist conclusion to the decision in *Privacy International*. After all, any expansion of the rule of law is surely a beneficial development, and the Supreme Court's advancement to access to judicial review of the IPT is likely to have a positive impact on digital rights. What is concerning, however, is how the judgement infringes upon the consensus established between the courts and Parliament on legislative ability to oust judicial review. As this paper has demonstrated, throughout the case law concerning ouster clauses, there has been a fine line drawn between the attempts to respect parliamentary sovereignty whilst simultaneously protecting the justiciable right of the courts in this jurisdiction. It has been, generally, a successful development.

In dialectic fashion, from a thesis in *Smith* via an antithesis in *ex parte Gilmore* to a hypothesis in *Anisminic* and the following case law, a consensus was met which held that, if an executive body or tribunal erred in a way that breached their jurisdictional capacity, they would not be exempt from review by the courts. As this paper has emphasised, however, this is a delicate consensus that the decision in *Privacy International* threatens.

In 1980, off the back of the decision in *Anisminic*, Wade celebrated that judges were beginning to stand up to illegal application of executive authority and discover "a deeper constitutional logic than the crude absolute of statutory omnipotence."[118] He was right: the law on ouster clauses is indicative of a hard-fought collaboration between the judiciary and Parliament that advanced key principles of the rule of law. Parliament remained sovereign, with the right to delegate its powers to executive tribunals, which would be protected provided they behaved legally. It is not coincidental that this developed alongside the doctrine of ultra vires. As Forsyth wrote in 1996, although the judiciary had "little concrete guidance as to the reach of judicial review and the scope and content of the various grounds of review… The judges no longer challenge legislative supremacy,

[117] *ibid* 144 (Lord Keith).
[118] Wade, *Constitutional Fundamentals* (n 15) 68.

the ouster clause remains attenuated, and there is sound constitutional basis for judicial review."[119] This description of the constitutional tenets is no longer as certain, because of how the decision in *Privacy International* has weakened the ouster clause, and legislative supremacy is faced off in the courts. The decision in this case has undermined the consensus between Parliament and the courts by removing the effectiveness of any potential ouster clause.

Qureshi, Tench and Hopkins noted that reference to Parliamentary intention is "strikingly absent" from the Supreme Court's judgement in *Privacy International*, which traditionally would have been the court's "touchstone" for judging an ouster clause.[120] In doing so, the Supreme Court has removed the intention of Parliament from the picture of justiciability. In this environment, what options lie before Parliament, if it wants to preclude the courts from review of tribunal determinations? As Edin has written, "The rule of law goes only so far as Parliament permits."[121] It is a fundamental principle of the rule of law that the judiciary are able to hold the executive to account, but if the courts show a willingness to step further out of the traditional bounds of the common law, into statutory intervention and invention through review, Parliament may feel the need to find other legal measures to oust its jurisdiction. "Abuse of legislative power," Winterton wrote, "should be rectified by political means, not by encouraging abuse of judicial power."[122] Disturbing the fragile consensus that was developed in *Anisminic* may prove to be dangerous, particularly at a time of constitutional uncertainty. There is a place in this dynamic, as the previous case law has demonstrated, for judicial review and parliamentary sovereignty to work in harmony. If manipulated as a process for introducing judge-made law, or undermining statutory authority, such a harmony is unlikely to last long.

VI. Conclusion

This paper, in exploring the historical and legal development of the ouster clause, has demonstrated the conceptual difficulty that ouster clauses pose. A question about an ouster clause is invariably a question about parliamentary

[119] Forsyth, "Of Fig Leaves and Fairy Tales" (n 105) 135–6.

[120] O Qureshi, D Tench and C Hopkins, "Case Comment: *R (on the application of Privacy International) v Investigatory Powers Tribunal and others* [2019] UKSC 22" (*UKSC Blog*, 23 May 2019) <http://ukscblog.com/case-comment-r-on-the-application-of-privacy-international-v-investigatory-powers-tribunal-and-others-2019-uksc-22/> (accessed 20 July 2019).

[121] Edin, "A Constitutional Right to Judicial Review" (n 14) 80.

[122] Winterton, "Parliamentary Sovereignty and the Judiciary" (n 108) 267.

sovereignty and the rule of law, framed within executive power and statutory interpretation. However, this complexity turns primarily on three main issues.

The first concerns the relationship between Parliament and the judiciary: ouster clauses are a fascinating depiction of the centuries' old conflict between law and politics playing out in real time. The courts' willingness to find in favour of ouster clauses mirrors judicial concerns that go much further than adherence to one single legal principle.

The second concerns the relationship between Parliament and the executive bodies, lower courts, and tribunals that seek to be protected by ouster clauses: it goes to the conceptual heart of what a body's legal jurisdiction consists of if it has been granted authority by Parliament, and which errors might undermine this. Despite the attempts by judges to move away from the traditional distinction between jurisdictional and non-jurisdictional errors, this division still haunts the case law, and will do so due to its fundamental relationship with the doctrine of ultra vires.

The third concerns the nature of legislative language: what, if any, words in statute will suffice to effectively oust the jurisdiction of the courts? And what, if any, interpretive action can the courts take to mediate this?

Underpinning all of these issues, this paper hopes to have demonstrated, are the practical political concerns that the ouster clauses touch. Due to their controversial nature, ouster clauses will only ever be used in politically-fragile situations where Parliament does not want judicial intervention. Ironically, it is this very aversion to review that draws the courts' attention. Should the courts stay away? This paper has argued that in certain cases, and for the sake of stability, yes. Sir John Laws rightly stated that, in the rule of law, "the greater the intrusion proposed by a body possessing public power over the citizen into an area where his fundamental rights are at stake, the greater must be the justification which the public authority must demonstrate."[123] This works more than one way for each of the branches of legal power, as the greater the intrusion proposed by the judiciary into an area of legislative jurisdiction, the greater must be the justification which the judiciary must demonstrate. With no written constitution, and to ensure a continued consensus between Parliament and the judiciary, the courts must not forget to rely on the principles of ultra vires to support judicial review of executive bodies. There is value in remembering the traditional distinction between errors, which provides certainty for all political and legal bodies.

The Supreme Court has created a clarity on the law of ouster clauses in its decision in *Privacy International*. There is an elegant simplicity in extending *Anisminic*

[123] Sir J Laws, "Is the High Court the Guardian of Fundamental Constitutional Rights?" (1993) PL 59, 69.

to hold any error as undermining jurisdiction, and it will be interesting to see where both Parliament and the case law go from this stage. If this is a dialectic, then the upheaval of *Privacy International* may well have started a new phase. It is likely that Parliament and the more conservative judiciary will develop logical ways around the direction of the Supreme Court.

As this paper has shown, however, this change to the definition of ouster clauses is not new. From *Smith* to *Anisminic*, and *Anisminic* to *Privacy International*, there is a long history of conceptual and linguistic changes which have led to new interactions and applications of administrative law. As Tucker has written, the story of the ouster clause is one of an "iterative game of 'cat and mouse'" between Parliament and the judiciary.[124] As concerns continue and evolve it is highly likely that this game will still play out, but within a different conceptual and linguistic arena. If a consensus between the courts and the legislature is to continue, particularly at a time when constitutional conventions are being questioned, rediscovered, and undermined, the ball is now in Parliament's court.

[124] Tucker, "Parliamentary Intention, Anisminic, and the Privacy International Case (Part One)" (n 82).

Cambridge Law Review (2019) Vol IV, Issue ii, 176–184

Vedanta Resources v Lungowe: A Pre-Existing Pocket of Negligence, or a Novel Scenario?

Tan Jin Hsi, Gabriel*

I. Introduction

The respondents, a group of 1,826 Zambian citizens living in the Chingola District in Zambia, had their health and farming activities damaged by alleged toxic emissions from the Nchanga Copper Mine ('the Mine') into watercourses on which they depend for drinking and irrigation. They brought claims in common law negligence and breach of statutory duty against Konkola Copper Mines ('KCM') and Vedanta Resources ('Vedanta'). Vedanata is the parent of a multinational group, incorporated and domiciled in the UK, and holds a significant majority stake of KCM and retains ultimate control of it.

The litigation largely concerns the jurisdiction of the courts of England and Wales to determine these claims against both defendants. The claimants rely upon article 4 of the Recast Brussels Regulation (Regulation (EU) 1215/2012 on Jurisdiction and the Recognition and Enforcement of Judgments in Civil and Commercial Matters) and para 3.1 of CPR Practice Direction 6B against Vedanta and KCM respectively.

In respect of the latter claim, as noted by the Supreme Court, the respondents had to demonstrate, *inter alia*, that there was 'between the claimant and defendant a real issue which it is reasonable for the court to try'.[1] The Supreme Court thus

* B.A. (Law) (Candidate), Girton College, University of Cambridge.

[1] CPR Practice Direction 6b, para 3.1.

had to decide if there was a real issue in respect of the duty of care claim against Vedanta.

II. Judgment

The judgments of the High Court and the Court of Appeal were largely the same. In both Courts, the point of departure for analysis was the *Caparo*[2] test.[3] In so doing, both Coulson J and Simon LJ appear to have been treating the case as a novel scenario. As clarified in *Robinson*,[4] novel categories of negligence should be developed incrementally, with the Caparo tripartite test used as a framework for inquiry as to whether such an incremental step should be taken. Both judges reached similar conclusions that a duty of care claim was arguable, and hence admissible.[5]

By contrast, in the Supreme Court, Lord Briggs, giving the unanimous judgment of the court, admonished against treating the present case as a novel scenario. Lord Briggs stated that: "the liability of parent companies in relation to the activities of their subsidiaries is not, of itself, a distinct category of liability in common law negligence… Everything depends on the extent to which, and the way in which, the parent availed itself of the opportunity to take over, intervene in, control, supervise or advise the management of the relevant operations… of the subsidiary."[6]

He continued at [54]: "Once it is recognised that, for these purposes, there is nothing special or conclusive about the bare parent/subsidiary relationship, it is apparent that the general principles which determine whether A owes a duty of care to C in respect of the harmful activities of B are not novel at all. They may easily be traced back as far as the decision of the House of Lords in *Dorset Yacht Co Ltd v Home Office* [1970] AC 1004".

It is trite law that following *Caparo*, the modern approach of the law of negligence operates on the basis of 'traditional categorisation of distinct and recognisable situations as guides to the existence, the scope and the limits of the varied duties of care which the law imposes'.[7] Lord Briggs' approach places the

2 *Caparo Industries plc v Dickman* [1990] UKHL 2.

3 *Lungowe v Vedanta Resources plc* [2016] EWHC 975 (TCC) [115]; [2018] 1 WLR 3575 (CA) [83].

4 *Robinson v Chief Constable of West Yorkshire* [2018] AC 736 (SC) [27], [29].

5 *Lungowe v Vedanta Resources* (TCC) (n 3) [121]; *Lungowe v Vedanta Resources* (CA) (n 3) [90].

6 *Vedanta Resources v Lungowe* [2019] UKSC 20 [49].

7 *Caparo* (n 2) 618; *Robinson* (n 4) [29].

case at bar into the same category of situations in which a duty of care is imposed as *Dorset Yacht*.

On the facts, Lord Briggs held that Vedanta assumed responsibility for the maintenance of proper standards of environmental control over the mining activities at the Mine. On this basis, it is well arguable that a sufficient level of intervention by Vedanta in the conduct of operations at the Mine may be demonstrable at trial, so that a duty of care may eventually be found.[8]

III. Analysis

Various commentators have discussed the jurisdiction element of the *Lungowe* litigation, but little has been said about Lord Briggs' approach to the duty of care issue.

This approach will be the subject of scrutiny in this note. Three issues fall to be discussed. Firstly, does the *Lungowe* scenario belong in the *Dorset Yacht* category of negligence? Secondly, and more broadly, is it advantageous for this line of cases—including *AAA v Unilever*[9] and *His Royal Highness Okpabi v Royal Dutch Shell Plc*[10]—to remain in the *Dorset Yacht* category? Thirdly, and consequently, would it be more appropriate for the *Lungowe* scenario to be treated as a novel (and separate) category of negligence?

A. Does the *Lungowe* scenario belong in the Dorset Yacht category of negligence?

The *Dorset Yacht* type of negligence can be broadly described as an exceptional situation wherein negligence liability is imposed on the basis of a 'pure omission'□exceptional because it contravenes the general principle that liability should not arise from purely omitting to do something.[11] The exception attaches liability to a person for the actions of a third-party which causes harm to the claimant. The lynchpin of this pocket of negligence, as was made clear in *Smith v Littlewoods*, is the special relationship between the defendant and the third-party, by virtue of which the former is responsible for controlling the latter.[12] This is exemplified by the facts of *Dorset Yacht*. Several borstal boys were working on an island under the control and supervision of three officers. One night, the officers left the boys to their own devices. The boys left the island and boarded, cast adrift,

8 *Vedanta Resources v Lungowe* (SC) (n 6) [61].

9 *AAA v Unilever* [2018] EWCA Civ 1532.

10 *His Royal Highness Okpabi v Royal Dutch Shell Plc* [2018] EWCA Civ 191.

11 *Smith v Littlewoods Organisation Ltd* [1987] AC 241 (HL) 247.

12 *ibid* 272; Michael A Jones, Anthony M Dugdale, Mark Simpson (eds), *Clerk & Lindsell on Torts* (22nd Ed, Sweet & Maxwell 2017) 8–55.

and damaged the plaintiff's yacht which was moored offshore. The Court held that because the officers had a responsibility to control the boys, they were liable for their misdemeanours, and thus owed the plaintiff a duty of care.[13]

At first blush, *Lungowe* appears to be an analogous case. As Lord Briggs rightly points out, parent and subsidiary are separate legal entities. The parent-subsidiary relationship does not attract liability per se, merely presenting an opportunity for control.[14] But one must take care to ask this question in a substantive manner—namely, whether the parent was controlling the aspect of the subsidiary which caused the harm. In short, the touchstone of the duty of care here is, like in *Dorset Yacht*, the control party A exercises over the actions of party B which gives rise to the harm.

However, the problem is that control can exist in many forms; it is a loose descriptor for many different types of special relationships between the defendant and the third-party in which it may be suitable to impose a duty of care. We have to be specific about the nature and extent of control. A recurring theme in the cases inhabiting the *Dorset Yacht* category of negligence is the custodial relationship between the defendant and the third-party, typified by physical custody of the latter.

In *Ellis v Home Office*,[15] the plaintiff, when a prisoner at Winchester Prison, suffered injuries as a result of an assault by another prisoner and sued the Home Office in negligence. In possessing physical custody of the third-party, the control the defendant had over him was immediate and absolute. A similarly high level of control can be found in *Carmarthenshire County Council v Lewis*,[16] wherein the third-party, a four-year-old boy, was attending a nursery school under the management of the appellant council. While not being attended by a teacher, the child ran into the road and caused an accident on the highway to a driver trying to avoid him. The child was in the physical custody of the school while he was in attendance, and the teachers undertook full responsibility for the safety and welfare of their students. In *Home Office*, the borstal boys, who caused damage to the plaintiff's yacht, were in the physical custody and supervision of borstal officers.

In fact, Lord Diplock's judgment in *Dorset Yacht* positions physical custody at the front and centre of the duty of care to be recognised. He held that:

> A is responsible for damage caused to the person or property of B by the tortious act of C (a person responsible in law for his own acts) where the relationship between A and C has the characteristics (1) that A has the *legal right to detain C in penal custody*

[13] *Dorset Yacht Co Ltd v Home Office* [1970] AC 1004 (HL) 1034, 1038, 1039, 1071.

[14] *Vedanta Resources v Lungowe* (SC) (n 6) [54].

[15] *Ellis v Home Office* [1953] 2 QB 135 (CA).

[16] *Carmarthenshire County Council v Lewis* [1955] AC 549 (HL).

> and to *control his acts while in custody*; (2) that A is *actually exercising his legal right of custody of C* at the time of C's tortious act and (3) that A if he had taken reasonable care in the exercise of his right of custody could have prevented C from doing the tortious act which caused damage to the person or property of B; and where also the relationship between A and B has the characteristics (4) that at the time of C's tortious act A has the *legal right to control the situation of B or his property as respects physical proximity* to C and (5) that A can reasonably foresee that B is likely to sustain damage to his person or property if A does not take reasonable care to prevent C from doing tortious acts of the kind which he did.[17] [Emphasis added]

Criteria (1), (2), (3) and (4) above strictly require some form of physical detention of the third-party by the defendant. Lord Diplock had also decided the case as a "rational extension of the relationship between the custodian and the person sustaining the damage which was accepted in *Ellis v Home Office*… and *D'Arcy v Prison Commissioners*… as giving rise to a duty of care on the part of the *custodian* to exercise reasonable care in *controlling* his detainee". [emphasis added].[18] It would surprise him today if he was told that his judgment was understood as only requiring the defendant to exercise some measure of control, loosely understood, over the third-party, absent any physical detention. The essence of this pocket of negligence, therefore, is physical custody, not merely control.

But even if we do not understand the *Dorset Yacht* line of cases so narrowly and stick with a 'control' in an open-ended sense, the pattern that emerges here cannot be ignored. Physical custody is a recurring theme in the cases inhabiting this pocket of negligence. At the very least, this underscores the high threshold of control necessary for an individual to be held liable for the actions of another.

Having regard to the evidence,[19] the most that can be said is that (1) Vedanta had overall oversight of KCM's activities, with a particular governance framework to prevent surface and ground water contamination by their operations, (2) Vedanta had a duty to provide various services and undertake feasibility studies into mining projects in accordance with accepted environmental practices, and (3) Vedanta was responsible for the provision of environmental and safety training to its subsidiaries. But Vedanta never had direct involvement with the Mine. 'Support', 'guidance', 'supervision' and 'oversight' (albeit to a high degree) are the words that come to mind, but it would be a stretch to say Vedanta 'controlled' the mining

[17] *Dorset Yacht* (n 13) 1063–4.
[18] *ibid* 1071.
[19] *Lungowe v Vedanta Resources* (CA) (n 3) [84].

operations which gave rise to the claims, at least not in the same sense exemplified by *Dorset Yacht*, *Carmarthenshire v Lewis*, and *Ellis v Home Office*.

By grouping the *Lungowe* case together with the *Dorset Yacht* line of cases, Lord Briggs compared the cases at a remarkably high level of abstraction. He, like many others, understood 'control' in a loose, open-ended sense. There was little regard for the specificity of control which the three cases featured, and, by extension, the high standard of control required. 'Control' becomes an umbrella term describing many different types of special relationships between defendants and third-parties, each of which warranting an exception to the pure omissions principle. It is not a precise term with which we can benchmark the duty of care in pure omissions cases, as it ideally should be.

It has to be recognised that *Lungowe* is, in an essential respect, different from other cases which co-exist in this sphere of negligence. At best, it exists on the periphery; at worst, it does not belong. Either way, it has to be acknowledged that recognising a duty of care in *Lungowe* would be an incremental expansion of the law. The only question is how big that incremental step is; this depends on how narrowly we construe the *Dorset Yacht* duty of care, and how different we think *Lungowe* is from that.

B. Is it advantageous for parent-subsidiary-claimant scenarios to remain in the *Dorset Yacht* pocket of negligence?

Here, we take a step back and look at the larger picture, including other cases featuring the same factual matrix as *Lungowe*. It would not be advantageous to analyse such cases through the prism of control supplied by *Dorset Yacht*, because it would be difficult for a duty of care to arise in such cases. As Sir Geoffrey Vos C mentioned in *Okpabi*, "it would be surprising if a parent company were to go to the trouble of establishing a network of overseas subsidiaries with their own management structures it if intended itself to assume responsibility for the operations of each of those subsidiaries. The corporate structure itself tends to militate against the requisite proximity".[20]

A tour d'horizon of the relevant cases supports this argument. In *Lungowe*, both the Court of Appeal and the Supreme Court found a duty of care claim to be arguable, but the facts of the case are exceptional. The parent company explicitly took charge of particular problems with discharges into water and the mine in Zambia, abiding to a governance framework to prevent such contamination. It undertook a contractual obligation to provide geographical and mining services to KCM. It was required to procure feasibility studies into mining projects in

[20] *His Royal Highness Okpabi* (n 10) [196].

accordance with accepted environmental standards. It provided environmental and technical training.[21]

Other cases remain a far cry from that. Rather than exercising effective control over the particular operation which gave rise to the claim, the parent companies in those cases merely exercise overall control of the subsidiary's operations. In *AAA v Unilever*, the claimants were workers and residents on a tea plantation in Kenya operated by the Kenyan operating company (UKTL), which was owned by a UK-registered parent company (Unilever). Following outbreaks of violence in 2007, the claimants claimed UKTL and Unilever had breached a duty of care to them in failing to take steps to prevent the violence, during which mobs killed, raped, and injured the appellants and their families. The parent company, Unilever, had "ultimate responsibility for the management, general affairs, direction and performance of the business as a whole",[22] but it was the subsidiary, UTKL, which prepared its own "crisis and emergency management" policy, and was not subject to the direction or advice of Unilever.

Okpabi tells the same tale. Following oil spills in Nigeria, claims were brought against the Nigerian operating company (SPDC) and its UK-registered parent (RDS), which was the parent of many subsidiaries worldwide (the Shell group). It was SDPC that was licensed in Nigeria to carry out the activities with which the spills were associated. The most that RDS did was issue mandatory policies, standards, practices, and a system of supervision and oversight across all its subsidiaries in the shell group, SPDC included,[23] but this was far from exercising material control.[24]

Following the recent surge of litigation involving parent companies and their subsidiaries—from *Chandler v Cape*[25] to *Lungowe*—multi-national corporations around the world would likely be scrambling to distance themselves as much as they reasonably can from the operations of their subsidiaries, so as to minimise the control they exercise and hence the liability they would incur should things go pear-shaped. This would make it even harder than it already is for a duty of care,

21 *Lungowe v Vedanta Resources* (CA) (n 3) [84].
22 *AAA v Unilever* (n 9) [17].
23 *His Royal Highness Okpabi* (n 10) [86].
24 *ibid* [122]–[123].
25 *Chandler v Cape* [2012] 1 WLR 3111; [2012] EWCA Civ 525.

on the basis of control, to be found moving forward. Should we continue down the lane of control, the prospect of redress for such victims looks bleak.

C. Would it be more appropriate for *Lungowe* to exist in a novel (and separate) category of negligence?

Various difficulties with Lord Briggs' approach in *Lungowe* have been demonstrated. Perhaps this explains why in *Thompson v Renwick*,[26] *Lungowe v Vedanta Resources* (in the Court of Appeal), and *Okpabi*, the Caparo factors were used, treating the case like a novel scenario.[27] It is suggested that, in the Supreme Court, *Lungowe* should have been treated the same.

One significant advantage to such an approach is the opportunity for the court to finally discuss the policy implications of recognising a duty of care in this area of law. When new cases fall within established categories of negligence, discussion about policy considerations—the fairness, justness, and reasonableness of finding a duty of care—become otiose, as clarified in *Robinson*.[28] And so it was not surprising for the UK Supreme Court to make no mention of the policy implications of finding a duty of care owed by Vedanta to the claimant.

But this area of law is a minefield of policy arguments. The issue of opening the floodgates to indeterminate liability, for instance, rears its ugly head yet again, especially if companies are, like the parent company in *Okpabi*, establishing mandatory policies and standards across all their subsidiaries across the world.[29] If a duty of care can be found in respect of the activities of one subsidiary, on that same basis, duties of care can potentially be found in respect of the activities of all its subsidiaries. There are also glaringly obvious implications this will have on the corporate structures of multi-national companies across the world. Some discussion about the economic advantages and disadvantages to finding the *Lungowe* duty of care is warranted.

Moreover, treating *Lungowe* separately from *Dorset Yacht* would achieve more rigour in the categorisation of situations in which a duty of care is found in this area of the law. On an abstract level, control is, like in Home Office, the kernel of the duty of care here. But control comes in many forms. It would not be splitting hairs to see that the control between the borstal boys and the officers is of a different nature from the control between Vedanta and KCM, the former relationship

[26] *Thompson v Renwick* [2014] EWCA Civ 635.

[27] *Thompson v Renwick* (n 26) [28]; *Lungowe v Vedanta Resources* (CA) (n 3) [83]; *His Royal Highness Okpabi* (n 10) [84]–[85].

[28] *Robinson* (n 4) [26].

[29] *His Royal Highness Okpabi* (n 10) [121].

defined by physical custody and close monitoring, the latter relationship defined by technical assistance and corporate responsibility over operations.

In this light, the new *Lungowe* category of negligence can still be broadly defined in terms of control, but in carving this new category the courts must be precise about what type of control they are looking out for. Helpful parameters can be found in *AAA v Unilever*, where Sales LJ, outlined two scenarios in which a duty of care would arise in this line of situations. Firstly, where the parent has in substance taken over the management of the relevant activity of the subsidiary in place of, or jointly with, the subsidiary's own management, or secondly, where the parent has given relevant advice to the subsidiary about how it should manage a particular risk.[30]

On the other side, the courts should be more specific in defining the parameters of the *Dorset Yacht* pocket of negligence. Rather than loosely grouping these cases as the 'control' cases, the courts can afford to be more precise about the nature of control which defines that category, perhaps by alluding to the presence of physical custody or the custodial relationship between defendant and third-party.

IV. Conclusion

There are evident difficulties with Lord Briggs' categorisation of *Lungowe* into the *Dorset Yacht* pocket of negligence. It has been suggested that the *Lungowe* scenario, and other cases like it, should form a separate category of negligence. This would give more definition to the parameters of these categories of negligence and provide their respective 'control' benchmarks greater precision. More importantly, it would give the courts an opportunity to discuss the policy implications of finding a duty of care.

Okpabi is on its way to the Supreme Court, and it is hoped that the court will make some amends at that point. But given the recency of the *Lungowe* judgment, which was unanimous, it is unlikely that much will change.

[30] *AAA v Unilever* (n 9) [37].

Cambridge Law Review (2019) Vol IV, Issue ii, 185–194

Personal Injury, Autonomy, and Johnson Matthey

Brian Ip*

I. Introduction

In *Dryden v Johnson Matthey plc*[1] (hereafter, *Johnson*), the Supreme Court pushed the concept of personal injury to the limits. Commentators believe that this was a pragmatic, policy-oriented response to compensate the Claimants for their financial loss by bringing such loss under a conception of personal injury. For this reason, they take issue with an obiter comment that is inconsistent with their reading of the case. However, this author believes that said comments are far from heretical. They reveal a different reading of the case, one that puts the Claimants' autonomy front and centre in a conception of personal injury. Though novel, this reading is one that can be supported and defended on the existing law.

II. Facts of the case

The Claimants were employed by the Defendants in positions that exposed them to platinum salts. Having been exposed to higher levels of platinum salts than that permitted by law, the Claimants developed a condition called sensitisation. Antibodies to the platinum salts were now present in their bloodstream. While asymptomatic, any further exposure to platinum salts would cause the Claimants to develop allergies with physical symptoms. But before the manifestation of any physical symptoms, the Claimants' sensitisation was detected through a routine skin prick test. In response, the Defendants either dismissed the Claimants entirely or redeployed them to other jobs that paid less. The Claimants brought proceedings

* BA (*Oxon*) (Candidate). I am grateful to Birke Håcker, Tom Krebs, Adam Perry, Alex Georgiou and Paul Fradley for their comments on earlier drafts. Special thanks also to Adam Walton and Katy Edgeley. Any errors that remain are my own.

1 *Dryden v Johnson Matthey plc* [2018] UKSC 18; [2018] 2 WLR 1109.

against the Defendants, alleging negligence on the part of the Defendants. They sought damages for loss of earnings resulting from their redeployment or termination.[2]

At first instance, Jay J held that platinum salt sensitisation was not itself a physical injury.[3] The Claimants thus had no cause of action. This was upheld in the Court of Appeal.[4] Sales LJ, giving the sole substantive judgment, treated the sensitisation as prior to and separate from any future allergy. His Lordship noted that there was a distinction between platinum salt sensitisation and the pneumoconiosis suffered by the Claimants in *Cartledge v E Jopling & Sons*.[5] Unlike pneumoconiosis, platinum sensitisation had no detrimental physical effects in the course of ordinary life.[6] It would only lead to future allergy (and so personal injury) if the Claimants continued to be exposed to platinum salts. As the Claimants were removed from such an environment, the sensitisation was now harmless, much akin to the symptomless and harmless pleural plaques in *Rothwell v Chemical & Insulating Co*.[7] For these reasons, his Lordship dismissed the Claimants' appeal.[8]

On appeal to the Supreme Court, the Claimants argued that platinum salt sensitisation, albeit asymptomatic, amounted to physical injury. The lost earnings resulting from the redeployment and dismissal were consequential economic losses that could be recovered by the plaintiff. The Claimants argued in the alternative that the Defendants had assumed responsibility to the Claimants in respect of their lost earnings.

On the other hand, the Defendants claimed that the negligent breach had merely increased the risk of physical injury, *i.e.*, the physical allergies, to the Claimants. The sensitisation was not in and of itself physical injury. Further, redeploying and dismissing the employees were steps taken to avoid the physical injury and so could not form the basis of any claim.

III. The judgment

Lady Black JSC, giving the only substantive judgment, rejected the Defendants' arguments. Lady Black held that platinum salt sensitisation could not be distinguished from platinum salt allergy, such that the sensitisation itself amounted to damage.[9] Although platinum salt sensitisation was asymptomatic,

[2] *ibid* [2]–[3].
[3] *Greenway v Johnson Matthey plc* [2014] EWHC 3957 (QB); [2015] PQIR P10.
[4] *Greenway v Johnson Matthey plc* [2016] EWCA Civ 408; [2016] 1 WLR 4487.
[5] *Cartledge v E Jopling & Sons* [1963] AC 758 (HL).
[6] *Greenway* (CA) (n 4) [21].
[7] *Rothwell v Chemical & Insulating Co* [2007] UKHL 39; [2008] 1 AC 281.
[8] *Greenway* (CA) (n 4) [30]–[33].
[9] *Johnson* (n 1) [37].

the presence of symptoms is not a prerequisite for personal injury. In *Cartledge*, the Claimants contracted pneumoconiosis in the course of their work. This condition caused scarring of the lung tissue. Although asymptomatic, the Claimants suffered a reduction in their lung capacity and were more susceptible to diseases.[10] The House of Lords held that the Claimants had suffered actionable physical injury, on the basis that in "unusual exertion or at the onslaught of disease he [*sic*] may suffer from his [*sic*] hidden impairment".[11]

Having put to bed any contention over the asymptomatic point, Lady Black proceeded to apply the decision of the House of Lords in *Rothwell*.[12] That case concerned workers who were exposed to asbestos fibres as a result of their employer's negligence. These workers later developed pleural plaques in their lungs and brought a claim against their employer for personal injury. The House of Lords dismissed their claim on the basis that they had not suffered any such physical injury. Damage, Lord Hoffmann said, was "an abstract conception of being worse off, physically or economically, so that compensation is an appropriate remedy".[13] The pleural plaques were simply a marker of exposure to asbestos dust. They did not cause any problems themselves and would not lead to anything harmful.[14] The present case was different, noted Lady Black. The Claimants started out with the bodily capacity to work around platinum salts. As a result of the Defendants' breach of duty, the Claimants lost this safety net and "therefore their capacity to work around platinum salts".[15]

This was a point of contention during the course of argument. Counsel for the Defendants argued that the Claimants had not been sensitised to something present in everyday life, but merely to a restricted and dangerous chemical only present in that particular employment context. As such, the risk of allergy could be averted simply by preventing them from working in situations involving platinum salts.[16] Recall that this too, was the view adopted by the Court of Appeal. But Lady Black disagreed, rejecting any such suggestion out of hand. In her words, "ordinary life is infinitely variable".[17] Indeed, there is nothing 'ordinary' in ordinary life – what is ordinary must be judged according to the particular individual. Lady Black gave the example of coffee tasters whose sense of taste was impaired, or expert perfumers who lost their sense of smell, such that they were no longer able to

10 *Cartledge* (n 5) 760–2.
11 *ibid* 779.
12 *Rothwell* (n 7).
13 *ibid* [7].
14 *Johnson* (n 1) [22].
15 *ibid* [37].
16 *ibid* [39].
17 *ibid* [39].

continue in that line of work.[18] Surely, they could be said to have suffered personal injury. Similarly, the sensitisation prevented the Claimants from working in a field of their choice. As a result, they earned less than they could have, which left them worse off. This was sufficient to amount to personal injury. For this reason, some commentators have alluded to this concept of "damage to the person" as opposed to mere "damage to the body".[19]

If Lady Black had stopped here, there would be relatively little controversy. However, Her Ladyship made some *obiter* comments that have troubled commentators.[20] Responding to questions by counsel, Her Ladyship held that the hypothetical worker who was about to retire or change occupation would nonetheless have suffered actionable personal injury.[21]

IV. Conceptions of loss and injury

A. Loss of future profits

At first glance, the approach adopted by the court in *Johnson* appears to be consistent with that adopted in professional negligence cases. These are commonly classified as 'loss of a chance' cases, but this terminology is misleading for they are not claims for a mere increase in risk. Green distinguishes the former from the latter in the following way. She says that in these cases "the chance exists independently of the breach of the duty, such that the breach affects the claimant's ability to avail herself of that chance, but not the substance of the chance itself".[22] In other words, the gist of the action is the lost opportunity to earn profits. There are three essential elements to such a claim:

1. The Claimant had the opportunity to make a choice about X.
2. But for the Defendants' breach, the Claimant would have acted differently in relation to X.
3. In acting differently, the Claimant could have availed themselves of a potential profit from X.

In order to prove that they had a shot at that potential profit, the Claimant has to show that they would have acted differently but for the Defendants'

[18] *ibid* [41].
[19] Robert Weir QC, 'What is a Personal Injury Anyway?' (2018) 4 Journal of Personal Injury Law 263, 269.
[20] Jonathan Morgan, "The Outer Limits of "Personal Injury"" (2018) 77 CLJ 461; Jarret J Huang, "*Dryden v Johnson Matthey*: The Boundaries of Actionable Damage" (2019) 82 MLR 737.
[21] *Johnson* (n 1) [45].
[22] Sarah Green, *Causation in Negligence* (Hart 2015) 154.

negligent action, to avail themselves of that chance.[23] Although Lady Black made no reference to this line of cases, her reasoning bears striking resemblance to that adopted in said cases. In *Johnson*, the Claimants had lost the capacity to work around platinum salts as a result of the Defendants' breach of duty. As a result, the Claimants were prevented from working in an occupation of their choice. Had they continued in their line of work they would have earned a higher salary. For this reason, Morgan considers that the "negative effects were essentially financial not physical ones".[24]

Understanding *Johnson* in these terms, we can see how Lady Black's obiter comments are problematic. Recall Lady Black said that the hypothetical Claimant who was about to retire or change jobs would nonetheless suffer personal injury. Such a Claimant would satisfy element 1 above. But he fails to satisfy element 2 or 3. The recent case of *Perry v Raleys Solicitors*[25] reminds us what we are to make of such a Claimant. The Claimant in *Perry* was afflicted with a particular condition as a result of his employer's negligence. He retained the Defendants as his solicitors to make a claim under a compensation scheme established by the Department for Trade and Industry. The Claimant alleged that the Defendants' negligence caused him to lose the chance to gain additional compensation. At trial, it transpired that the Claimant was unable to prove he met the criteria for additional compensation. The Supreme Court allowed the Defendants' appeal on this basis. It held that unless the Claimant could prove that but for the Defendants' breach he would have acted differently, the Claimant would not have lost any 'chance' at all. Accordingly, Lady Black's hypothetical Claimant would not have suffered personal injury. Commentators have seized on this reasoning to criticise this part of Lady Black's judgment. Huang argues that Lady Black conflates the issue of damage with that of quantification.[26] Any Claimant who was about to retire could only be said to have suffered negligible injury, for which he could not recover. Morgan too seems troubled by this. He raises the example of the hypothetical university lecturer who was negligently exposed to platinum salts.[27] It would be problematic, says Morgan, for such an individual to recover. Yet this would be precisely the result should Lady Black be correct.

The validity of these criticisms, however, depends on the above reading of *Johnson* being correct. Yet the court did not reason in quite those terms. At paragraph 44 of her judgment Lady Black says: "Once the sensitisation is

[23] *Perry v Raleys Solicitors* [2019] UKSC 5; [2019] 2 WLR 636 [19]–[23]; *Allied Maples Group Ltd v Simmons & Simmons* [1995] 1 WLR 1602 (CA); Green (n 22) 154.

[24] Morgan (n 20) 464.

[25] *Perry* (n 23).

[26] Huang (n 20) 742–3.

[27] Morgan (n 20) 463.

identified as an actionable injury in its own right, the company's arguments that the Claimants are, in reality, claiming only for their lost earnings and therefore for pure economic loss also falls away."

Lady Black juxtaposes actionable injury with lost earnings in rejecting the Defendants' arguments on pure economic loss. This contrast suggests that the court did not regard the lost earnings as *constitutive of*, but rather *consequential on* the Claimants suffering personal injury.[28] If this is correct, then personal injury must be established independently from and prior to the claim for lost earnings. Consider the hypothetical situation of a Claimant that has developed sensitisation to alcohol, such that the consumption of a single drop will cause him to break out in hives. Our Claimant's work does not involve alcohol of any kind, nor does he intend to work around it. In his personal life, he is a teetotal. Nevertheless, we would still say that he has suffered personal injury, even if he cannot be said to have lost any future earnings.[29] If so, then any explanation for why the sensitisation amounted to personal injury must lie elsewhere.

B. Rights to autonomy

Such an explanation might, counter-intuitively, be found in Lady Black's obiter comments. The Claimants in *Johnson* did not lose the opportunity to make a gain, but rather suffered an infringement of some limited right to make a decision about themselves. It is the infringement of this right that constitutes damage. Such a conception would involve:

1. A breach of the posterior duty (the Defendants' duty to ensure safe working conditions), the breach of which causes;
2. The infringement of an anterior right (the right to make future decisions about one's occupation)

This view of rights protecting rights is given support by the recent case of *Morris-Garner v One Step (Support) Ltd.*[30] That case dealt with the issue of when a claimant can recover 'negotiation damages' for breach of contract. Before this case, the courts were of the view that negotiation damages were a discretionary remedy for breach of contract.[31] But Lord Reed, with whom the majority agreed, rejected this line of thought. Negotiation damages were awarded for breach of contracts that created or protected ancillary rights valued as assets. One such example is a contractual license to use intellectual property. In such cases, the

[28] *Johnson* (n 1) [44].
[29] My thanks to Birke Håcker for this example.
[30] *Morris-Garner v One Step (Support) Ltd* [2018] UKSC 20; [2018] 2 WLR 1353.
[31] *ibid* [81], see also *Experience Hendrix LLC v PPX Enterprises Inc* [2003] EMLR 25 (CA) [34]–[35] (Mance LJ).

breach of contract would cause the loss of a valuable asset, to be measured by the economic value of the protected right.[32]

In the realm of 'loss of a chance', Green argues convincingly that the claimant in *Chester v Afshar*[33] recovered because she lost a chance to decide whether to undergo the surgery.[34] Dr Afshar was a surgeon who advised Ms Chester to undergo a spinal surgery. He failed to inform Ms Chester that the procedure carried a small risk that the patient would develop cauda equina syndrome. Ms Chester agreed to the surgery, and subsequently developed the syndrome. She sued the Defendants in negligence. On appeal to the House of Lords, their Lordships dismissed the Defendants' appeal and held Dr Afshar liable in negligence.[35] Yet Ms Chester did not lose the opportunity of gaining a better outcome. The risk of injury was inherent within the operation. At trial, she even conceded that had she been told about the risk of injury, she might still have agreed to the surgery at some other time.[36] What she suffered was an infringement of her right to make informed decisions concerning bodily integrity,[37] which in itself amounted to damage.

Steel astutely observes the difference between this category of cases and that of 'lost opportunity to make a profit' cases in his analysis of *Chaplin v Hicks*.[38] The Claimant was a beauty contestant who was shortlisted for an interview. On the basis of the interview, a select few were to be awarded actress engagements. The Defendants failed to inform the Claimant of her interview time, thereby breaching his contract.[39] On appeal, the Court of Appeal allowed the Claimant's appeal for substantial damages. Yet their Lordships did not do so on the basis that she had lost a chance of winning. Rather, the Claimant was entitled to recover because she had lost the right to participate in the contest. As this was a right of value, the Claimant was able to recover damages amounting to the pecuniary value of that right.[40]

The courts have adopted a similar approach in cases involving wrongful conception. In *Rees v Darlington Memorial NHS Trust*,[41] the Claimant was a disabled mother who went for a sterilisation procedure at the Defendants' hospital. The procedure was negligently performed by the Defendants. The Claimant later successfully conceived a healthy child and brought a claim against the Defendants

32 *One Step* (n 30) [92]–[95].
33 *Chester v Afshar* [2005] 1 AC 134 (HL).
34 Green (n 22) 158-160.
35 *Chester* (n 33) [11].
36 *ibid* [31]–[32].
37 *ibid* [16]–[18], [24]. This right was recognised in *Montgomery v Lanarkshire Health Board* [2015] UKSC 11; [2015] 1 AC 1430 albeit not in this context of rights protecting rights.
38 Sandy Steel, *Proof of Causation in Tort Law* (CUP 2015) 296.
39 *Chaplin v Hicks* [1911] 2 KB 786 (CA), 786–8.
40 *ibid* 793, 796.
41 *Rees v Darlington Memorial NHS Trust* [2004] 1 AC 309 (HL).

for damages. The House of Lords rejected the claim, on the basis that the birth of a child is not a wrong. But their Lordships nevertheless allowed the Claimant to recover for the lost autonomy to order her family life.[42]

This is consistent with the approach adopted by the court in *Johnson*. The Defendants breached their duty of care to provide a safe working environment for the Claimants. As a result, the Claimants were exposed—and developed sensitisation—to platinum salts. This sensitisation amounted to an infringement of the Claimants' limited right to choose their occupation, the economic value of which was measured by the possibility of continuing in the same line of work.

V. When such rights might exist

If this is correct, the question that follows is: 'when might such rights exist?'. It is clear that we have no general free-standing right to autonomy. Almost every action has the potential to affect the autonomy of others. If the law treats a mere reduction in valuable choices as actionable damage, it would open the door to large and indeterminate liability in tort.[43] As Steel notes, this parallel is reflected in the approach the law takes in dealing with pure economic loss.[44] Losing wealth entails losing valuable options that could be pursued with that wealth. But no one seriously suggests that we have a freestanding right to economic gain.[45]

That is not to say, however, that we should not recognise a limited duty in some circumstances. As Cane notes, whether A is said to have caused something to happen to B depends in part on B's obligations to A.[46] A plausible explanation is that in exceptional relationships of trust and confidence, the law deems B to have assumed responsibility to A not to negligently impinge on his autonomy.[47] Such relationships would include employer and employee (*Johnson*) or doctor and patient (*Chester*, *Rees*). Already there is some support for this in the law. The courts have recognised an implied obligation of mutual trust and confidence in employment contracts, as a response to the disparity of power between the employer and

[42] Craig Purshouse, "Judicial Reasoning and the Concept of Damage: Rethinking Medical Negligence Cases" (2015) 15 Medical Law International 155, 166–7; Donal Nolan, "New Forms of Damage in Negligence" (2007) 70 MLR 59, 79.

[43] Jane Stapleton, "Duty of Care and Economic Loss: A Wider Agenda" (1991) 107 LQR 249, 254–8; Leonard Hoffmann, "Causation" (2005) 121 LQR 592, 600.

[44] Steel (n 38) 349.

[45] Robert Stevens, *Torts and Rights* (OUP 2007) 26.

[46] Peter Cane, *Responsibility in Law and Morality* (Hart 2002) 132; see also HLA Hart and Tony Honore, *Causation in the Law* (2nd Ed, Clarendon Press 1985) 33; Clarke, *et al*, "Causation, Norms, and Omissions: A Study of Causal Judgments" (2015) 28 Philosophical Psychology 279.

[47] Steel (n 38) 349; Morgan (n 20) raises the point about assumption of responsibility, but in relation to pure economic loss.

employee.[48] If the employee suffered from reduced employment prospects as a result of the breach, then the employer will be made liable for any continuing financial loss.[49] In Lord Nicholls' words, "employers must take care not to damage their employees' future employment prospects, by harsh and oppressive behaviour or by any other form of conduct which is unacceptable today as falling below the standards set by the implied trust and confidence term".[50]

Understanding damage as the infringement of an ancillary right helps to address Morgan's concerns. Whether his hypothetical university lecturer has suffered damage depends on the scope of his employer's duty to him.[51] It is at least arguable that a university's duty does not extend to protecting a lecturer's choice to work around platinum salts. We can rest easy at night knowing *Johnson* will not lead to large, indeterminate liability.

This conception is supported by a close reading of Lady Black's speech in *Johnson*:

> The restrictions on the work that can be done by Claimants… are attributable to the sensitisation, to which the protective provisions of the collective agreement *were a response*. These provisions reflect the fact that, because of the negligence… these Claimants' bodies are now in such a state that they need to avoid further exposure to platinum salts… But the *need for sensitised individuals to avoid exposure would apply whether or not there was a collective agreement*… whether the employer was Johnson Matthey or another employer who imposed no comparable restrictions.[52] [Emphasis added]

Huang notes that, strictly speaking, the Claimants could still work around platinum salts.[53] They did in fact do so. At trial, it was found that the Claimants were in fact sensitised before such sensitisation was detected. Yet they remained employed in positions that exposed them to platinum salts.[54] Any right that was infringed has its contents prescribed as a matter of law - the court deems the employer to have assumed responsibility to the Claimant, such that the employer came under a duty to avoid impinging on the Claimant's career opportunities.

[48] *Malik v Bank of Credit and Commerce International SA* [1998] AC 20 (HL).
[49] *ibid*; see also *Abbey National plc v Chagger* [2009] EWCA Civ 1202; [2010] ICR 397.
[50] *ibid* 38.
[51] *C.f.* the approach adopted in *Palsgraf v Long Island Railway Co* (1928) 248 NY 339, and *South Australia Asset Management Corp v York Montague Ltd* [1997] AC 191 (HL).
[52] *Johnson* (n 1) [43].
[53] Huang (n 20).
[54] *Johnson* (n 1) [43].

Future courts will have to address questions such as when this anterior right will exist, and what will amount to an infringement of this right.

VI. Conclusion

Dryden v Johnson Matthey presents an opportunity for us to reconsider our conception of what amounts to personal injury. While some believe that the case is authority for pecuniary loss as personal injury, the better reading of the case is that it recognises people can be made worse off by the loss of valuable choices.